

## Journal of the Gypsy Lore Society

Edinburgh : The Society, 1889-1999

<https://hdl.handle.net/2027/inu.30000071184661>

# HathiTrust



## [www.hathitrust.org](http://www.hathitrust.org)

**Creative Commons Attribution-NonCommercial-NoDerivatives**

[http://www.hathitrust.org/access\\_use#cc-by-nc-nd-3.0](http://www.hathitrust.org/access_use#cc-by-nc-nd-3.0)

This work is protected by copyright law (which includes certain exceptions to the rights of the copyright holder that users may make, such as fair use where applicable under U.S. law), but made available under a Creative Commons Attribution-NonCommercial-NoDerivatives license. You must attribute this work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work). Only verbatim copies of this work may be made, distributed, displayed, and performed, not derivative works based upon it. Copies that are made may only be used for non-commercial purposes. Please check the terms of the specific Creative Commons license as indicated at the item level. For details, see the full license deed at <http://creativecommons.org/licenses/by-nc-nd/3.0/>.

# **The Speech of the Polska Roma: Some Highlighted Features and Their Implications for Romani Dialectology**

**Yaron Matras**

---

*Selected features of the dialect of the Polska Roma are considered, based on the Lešaki variety. Despite lexical, phonological, and some morpho-phonological evidence for an historical affinity with a "Northern" branch of Romani, the majority of the salient morpho-syntactic characteristics can be explained as the outcome of areal diffusion among Romani dialects, or as preferences deriving from underlying variation in Romani as a whole. This suggests that a comprehensive dialectological investigation of Romani should take a layered view, distinguishing inherited (genetic) divisions from acquired (areal-convergent) traits. In the discussion attention is given to some points of interest for a morpho-syntactic typology of Romani dialects.*

---

## **1. The Northern Dialects of Romani**

Among the dialect groups posited in recent classifications of Romani varieties (see Bakker and Matras 1997; also Boretzky and Igla 1991: 8–10; Kaufman 1979), the so-called Northern branch appears to be the most diversified. Bakker (1999) identifies the Northern branch as comprising a) the dialects of the Finnish, Baltic, North Russian, and Polish Roma; b) those of the Sinti-Manuš group in Germany, the Netherlands, France and out-migrants in Italy, former Yugoslavia, and elsewhere; c) those of southern Italy (Calabria and Abruzzi); as well as d) the now extinct varieties of British, Iberian, and Scandinavian Romani. From the geographical dispersion it is clear that the term "Northern" is a rather arbitrary label coined primarily for reference purposes, inspired by the predominance of the varieties in the north of Europe. It does not however imply either current confinement to northern Europe or original formation and subsequent outward spread from there.

---

Yaron Matras is Research Fellow in Linguistics at the University of Manchester, Oxford Road, Manchester M13 9PL, UK; e-mail y.matras@man.ac.uk.

Indeed, the exact nature of the historical connections among the Northern dialects and the question of whether they reflect an underlying subdivision within early or Byzantine Romani are likely to remain unanswered as long as historical reconstruction continues to rely entirely on contemporary linguistic evidence. Non-linguistic sources merely portray a movement beginning at the end of the fourteenth century of small groups of Roma, usually referred to in the documents as “bands of Gypsies” and consisting of extended families or clans of several dozen persons each, through central Europe and outwards in all directions but to the southeast, whence they appear to have originated. Some of those could have been the ancestors of speakers of Romani dialects now dispersed, for which Miklosich (1872–1880) had already reconstructed shared migration routes based on the various layers of lexical borrowings.

A significant part of the linguistic data on which modern postulations of a Northern branch rest derives from the so-called Para-Romani varieties, better defined as Romani-derived vocabularies used mainly by itinerant tradespeople of either Romani or, more probably, of mixed ethnic origin. These vocabularies are embedded into non-standard varieties of the various majority languages, e.g. Spanish, English, Scandinavian, or German (see. Bakker and Van der Voort 1991; Boretzky and Iglá 1994; Matras 1998c). Their use for the purpose of classifying underlying dialects of Romani presupposes a direct diffusion of lexical items from Romani varieties which had once been spoken in the community or were otherwise coterritorial with the emerging Para-Romani vocabularies. Thus, Iberian Para-Romani or Caló is assumed to have descended from an Iberian dialect of Romani which it replaced (see Bakker 1995), Scandinavian Para-Romani is supposed to have succeeded a variety akin to Finnish Romani (cf. Ladefoged 1998), etc. Only for British Romani, however, is there extensive documentation of such coterritorial Romani varieties from the time prior to their becoming extinct (e.g. Smart and Crofton 1875; Sampson 1926).

It is in principle possible that Para-Romani vocabularies did not enter non-standard usages from a coterritorial dialect of Romani, but that they were transmitted through diffusion of specialized vocabularies or trade jargons. Such “traveling” lexicons were once widespread in Europe, as can be seen from the close affinity between numerous items in secret on-the-road vocabularies and various slangs deriving from those which include Romani elements. German Manisch, for example, the secret jargon of itinerants in the city of Giessen, has been shown to have acquired part of its Romani component from other Rotwelsch and Jenisch special vocabularies (Lerch 1976). Disregarding such possible complications, the reconstruction of a broader Northern branch of Romani remains, due to the prominence of Para-Romani vocabularies, largely confined to lexical features and the morpho-phonological properties they represent.

Some of the features listed by Bakker as Northern “forms” or “structures,” however, are better regarded as “strategies” which the respective varieties share. These strategies include the widespread tendency to use group-specific autononyms (*sinti*, *romaničel*, *manuš*, *kale*) rather than the term *roma*. The latter, in its singular forms *m. rom f. romni*, is only used for ‘husband/wife’, though the language is called *romanes*. In addition, some Northern dialects show internal lexical innovation through genitive derivations and compositions (*Sinti pimaskeri* ‘cigarette’, < *pi-* ‘drink’; *rakerpaskro čiruklo* lit. ‘talking bird’ for ‘parrot’, etc.), as well as widespread usage of cryptic place names (e.g. *Sinti xačardo them* lit. ‘burned country’, for *Brandenburg*; on cryptolalic usages in Northern dialects see also Günther 1915: 16–19, as well as Wagner 1937).<sup>1</sup>

Although classified as overall “Northern features” by Bakker, it appears to me that these strategies are rather typical of a *subtype* of dialects within the Northern branch, situated in the westernmost regions (and extending via Sweden also to Finland). Characterizing this subtype are extra-linguistic factors that are reflected in language: the isolation of the groups of speakers; the fact that until fairly recently, and partly to this day, they were typically peripatetics engaged in itinerant occupations; their strict cultural conservatism; and the extreme persecution to which they were subjected from the sixteenth century onwards. All this helps explain their strong immediate and group-specific, rather than more global (“Romani”) ethnic orientation as expressed in their diverse autononyms; their employment of language for secretive purposes, hence the frequent reliance on internal innovations rather than on lexical borrowings from the surrounding contact languages; and their tendency to mix with, and partly assimilate into other itinerant groups, hence the tendency among some Northern dialects to give rise to, and ultimately merge with or be replaced by Para-Romani vocabularies.

Conversely, varieties of the Northern branch whose speakers settled closer to the comparatively dense Romani population centers in central Europe have become integrated into a geographical continuum of Romani dialects. This dialect continuum shows a spread of isoglosses resembling, to an extent at least, that of “territorial” languages. In this respect the Sinti varieties of Germany and neighboring regions could be considered as “fringe” or even “frontier” dialects. Their speakers are more remote from the Romani population centers, and as a very small minority they are ethnically isolated; they have had considerable contact with non-Romani itinerants and have influenced the emergence of itinerant cultural identity, particularly Jenisch identity, including the secret vocabularies used by the Jenisch. Yet they have resisted linguistic assimilation and retain their inflected dialects of Romani, which nevertheless differ considerably from the “average” Romani of central or southeastern Europe.

The dialect of the Polska Roma could be said to constitute a linguistic sub-fringe, being the westernmost of the Romani dialects under Slavic influence while at the same time located at the edge of the plane of territorial Romani stretching from the Balkans to central Europe. One superficial though significant isogloss which appears to separate Polish from German dialects of Romani is the use of the autonym *rom*, found among the *bergitka roma* in the south, the *polska roma* in central Poland, and among the *sasitka roma* or “German Roma,” whose dialect is a subdivision of that of the German Sinti. The use of *rom* extends into North Russian and Baltic Romani as well, while group-specific autononyms prevail to the west (and, via Sweden, in Finland).

Of the inflected Romani dialects that form part of the Northern group, full grammatical descriptions cover Sinti (Liebich 1863, Finck 1903, Holzinger 1993), Welsh Romani (Sampson 1926), Finnish Romani, and the North Russian dialect of the *xaladiitka roma* (Sergieviski 1931; Wentzel 1980). Less accessible are the Romani dialects of central Poland. As in other countries, there are of course several different groups of Roma who have settled in Poland. They include the Sinti (*sasitka roma* ‘German Roma’), speakers of the Carpathian Central dialect (*bergitka roma*, ‘Carpathian’ or ‘mountain Roma’), speakers of North Russian dialects, and Lovara, who speak a Vlax (Romanian-Transylvanian) variety.

The term *polska roma* is reserved for the Romani population with well-established roots in the central districts of the country. In social organization they stand out in their allegiance to a *šero rom* or ‘head Rom’ who exercises authority in matters relating to community affairs and interfamily relations (Ficowski 1950). Unlike the Carpathian Roma of southern Poland or *bergitka roma*, the *polska roma* are considered to have been itinerant until fairly recently (cf. Bartosz 1994: 17). Their dialect is closely related to that of the *xaladiitka roma* or “North Russian Roma” as described by Wentzel (1980) and Sergieviski (1931), as well as to other Baltic Romani dialects (e.g. the Lithuanian dialect reported on by Dowojno-Sylwestrowicz 1889). It also shares some features with the Carpathian (Rumungri,<sup>2</sup> or Central) dialect of the *bergitka roma* of southern Poland, as documented by Kopernicki (1930) and Rozwadowski (1936).

The Polska Roma dialect might thus be expected to form a geographical link between a subdivision of the Northern branch and the geographically adjacent continuum of dialects known as the Central branch. The arrangement of isoglosses in the Romani dialects of this area has received intensified attention recently<sup>3</sup>, and it is for this reason that I choose to present the following sketch of highlighted features in one of the Polska Roma varieties, although the available documentation can by no means allow for a comprehensive and detailed description at this stage. I confine myself in the following to some observations based on the particular sub-variety of the *lešake roma*.<sup>4</sup> The speakers considered have been settled for quite

some time in Pabianice, south of Łódź. They appear to have derived their name from that of the Leszno district, or, some claim, the "Lesz forest."

Readers with no structural familiarity with Romani are referred to the grammar of the closely related dialect of the Xaladitka Roma by Wentzel (1980); to the grammars by Holzinger (1993) for the more remotely related Northern dialect of the Sinti, and by Sampson (1926) for the Welsh dialect, which is also part of the Northern branch; finally, to the grammatical sketch of the geographically proximate Central dialect of Slovakia in the dictionary by Hübschmannová et al. (1991).

## 2. Lexicon

Perhaps the most outstanding structural property (i.e., as opposed to "strategy," see above) shared by the Northern dialects is the presence of a series of lexical items not found in the other branches. Northern words in the Polska Roma dialect include *bolipen* 'sky', *čhiben* 'bed', *čiro* 'time', *dudali* 'window', *bar* 'thousand', *gušt* 'finger', *rakirel* 'to speak', *kuty* 'a little'.

A further set of Northern lexical items is shared with the southern Balkan dialects. Some examples are *her* 'foot', *taša* in the meaning 'tomorrow', *čhorja* 'beard', *men* 'neck', *šuvlo* 'fat', *rikirel* 'to hold', *kisyk* 'pocket', and the deictic expressions *dava* 'this', *(a)daj* 'here'. An additional class of words is in principle shared with the other dialect branches with the exception of Vlax (Romanian-Transylvanian Romani). Examples are *rukh* 'tree', *džipen* 'life'. Words common in other branches of Romani but with a distinct meaning in the Northern dialects include *kirmo* 'reptile' (elsewhere 'worm'), *naj* 'fingernail' (elsewhere 'finger'), and *taša* in its double meaning, 'tomorrow' and 'yesterday' (in the Balkans 'tomorrow' only).

The early migration of the Polska Roma from German-speaking territory is reflected in the retention of a broad German-derived vocabulary. Integration patterns, including gender assignment, which often differs from the underlying gender in German, and the athematic endings<sup>5</sup> attached to the nouns (-a for feminine, -o for masculine), are shared with German Romani (Sinti). Some examples are *štała* 'cellar', *korba* 'basket', *tyša* 'table', *špera* 'trace', *zanta* 'sand', *bilta* 'picture', *frajda* 'joy', *moto* 'instant, time', *fraj* 'free', *malinet* 'to paint', *frajdžuvel* 'to rejoice'. Many of these, along with the same adaptation patterns, are found in North Russian and Baltic Romani dialects as well, indicating shared origin and migration patterns. It is noteworthy that no grammatical items of German origin are encountered. This could suggest rather superficial contact with German, or possibly that the intensity of subsequent influence from Polish has led to the replacement of grammatical items which may have been borrowed earlier through items from the contemporary contact language (for a discussion of borrowing layers see Matras

1996). Naturally, a great number of Polish elements has been incorporated into the dialect. Owing to the position of Polish as a second community language (and the dominant language outside the community), virtually any item from Polish is a potential component of the dialect.

### 3. Phonology

On the whole Romani dialects share relatively few phonological features. Those include the presence of distinctive aspiration with voiceless stops/affricates and the tendency towards a reduced, Balkan-type vowel system consisting of five cardinal vowels and occasionally a peripheral centralized vowel.<sup>6</sup> While the first feature (distinctive aspiration) is clearly inherited from Indic, the second (Balkan-type vowel system) can be considered an outcome of the process of Balkanization following settlement in Europe and Asia Minor in the early Romani period, sometime between the tenth and fourteenth centuries (see Matras 1994: 13–19). With other phonological features, convergence with the dominant contact language after three to four generations of contact is common. Indeed, some Romani dialects also show changes to the vowel system and even a reduction of distinctive aspiration. Rather than take an inventory of phonemes, the discussion here will therefore focus on features that are an outcome of historical developments affecting particular phonemic sequences, and which we are able to compare with other dialects of the language.

Four areas of historical phonology deserve particular attention: 1) features typical of the Northern branch, 2) those shared with northeastern varieties of the Northern branch and with that of the neighboring Xaladitka Roma in particular, 3) those that are in principle northeastern features but constitute a further development in the variety under consideration, and finally 4) unique characteristics of the Polska Roma dialect.

#### 3.1. General Northern Features

Initial *v* in *vavir* ‘other’.

Retention of *t* (as opposed to palatalization or affricatization) in *xytlet* ‘to catch’, *tykno* ‘small’, *tiro* ‘your’, *buty* ‘work’; this is partly shared with southern Balkan dialects.

Retention of *k* (in place of palatalization) in *kšil* ‘butter’, *kirat* ‘cheese’.

Simplification of underlying *nd* in *maro* ‘bread’, *jaro* ‘egg’ (compare Central dialects *maro* but *jandro*), and retention of *n* in *pani* ‘water’; both are generally shared with southern Balkan varieties.

Root vowels *u* in *mukhet* 'to leave' (compare Vlach *mekh-*), *a* in *daj* 'mother' and *čhaj* 'girl' (compare Vlach *dej, šej*), and *i* in *džinet* 'to know', and the possessives *miro, tiro* 'my, your', genitive case *-kiro*.

An *s*-less form in *grai* 'horse' (Vlach *gras(t)*).

No *v* in *saro, sare* 'all' (Vlach *savoro*) but retention of *v* in *čovahani* 'witch', *dyves* 'day'. Here too there are parallels with southern Balkan.

Retention, as in southern Balkan, of "long" genitive forms in *-kiro* etc. (Vlach *-ko* etc.).

### 3.2. Features of the Northeastern or Polish - North Russian - Baltic Group (Shared with the Xaladitka Dialect)

Loss of *s* in the 3rd person singular of the past tense of verbs, e.g. *javja* 'came', *lija* 'took', *geja* 'went'. This is a common development in Romani and is found within Vlach (Kelderash, in contrast with Lovari), within the Central dialects, as well as in the southern Balkan varieties. It contrasts nevertheless with the neighboring dialects of southern Poland and with Sinti, where *s* is preserved (*javjas* etc.).

*er* > *ir* in *vavir* 'other', and in *-kirel*, a verb-derivational suffix, from *kerel* 'to do', e.g. *rakirel* 'to speak'.

*Cr* > *Cš* in *pšat* 'brother'.

### 3.3. Further Development of Northeastern Features Generally Shared with Xaladitka

*V* > *jV*. This is shared with other Northern as well as Central dialects for the third person pronouns *jov* 'he', *joj* 'she', *jone* 'they'. It is shared with other Northern dialects for *jaro* 'egg'. For Xaladitka, Wentzel (1980) has *javel* alongside *avel* 'to come', and Sergievski (1931) has *javne* 'they came' but *avjom* 'I came'; the Polska Roma dialect has consistently *javel*, *javne*, *javjom*, and in addition *jame* 'we' (Wentzel: *ame*), and *jandel* 'to bring' (Wentzel: *anel*).<sup>7</sup>

*ajn, an, on* > *en*. We find this in Xaladitka for *čergen* 'star', here *čerhen*; in the Polska Roma dialect also in *vurden* 'car', *moxten* 'box'.

Centralization of *a* > *y*. A parallel development is found in Sinti, where *a* may shift to *e* or Schwa. As in Xaladitka, we find *syr* 'how', *-tyr* the ablative suffix, *sys* 'was', *maškyr* 'between', but in the Polska Roma dialect this extends to include the remoteness tense marker used to form the imperfect, the past tense of the copula, and the pluperfect/conditional: *-ys*. In addition, *o* > *y* is occasionally observed, as in *kotyra* 'pieces'.



Centralization of *i* > *y* is shared with Xaladitka in *sy*, 'is', *dykhet* 'to see', *tykno* 'small', etc., but is extended to include further items, e.g. *ryg* 'side', as well as, irregularly however, the feminine adjectival ending *tykny* alongside *tykni* 'small'.

No initial *a*. General variation among Romani dialects in this area evidently derives in part at least from a stem alternation already found in Middle Indo-Aryan. On the continuum of Romani dialects that tend to show initial *a*, the Polska Roma figures on the minimal side. Along with Xaladitka it has *rakhet* 'to find', *kharel* 'to call', *šaref* 'to praise', but also *kana* 'now' (Xaladitka *akana*), *čhel* 'to stay' (Xaladitka *ačhel*), *maro* 'our' alongside *jamaro* (Xaladitka *amaro*).

### 3.4. Particular Features

The velarization of underlying aspirated phonemes resulting in the realizations [px, tx, kx, čx] is shared with other Romani dialects in contact with Polish, Ukrainian, Russian, and the Baltic languages. The variety under consideration, however, preserves the phonemic distinction between /h/:/x/. Insertion of initial aspiration Ø > *h* is even observable, e.g. *hučo* 'high' (Xaladitka *učo*).

Palatalization of consonants is common, as in other Romani varieties in contact with Slavic languages, but its phonological status is not sufficiently clear at this stage. Hence it is not represented in the notation used here, and I do not distinguish palatalization of dentals from affricatization: *vušt*, pl. *vušča* 'lips' (= *vušt' a*), *gušt* pl. *gušča* 'finger' (= *gušt' a*), (but compare *kašt* pl. *kašča* 'wood'); the latter development stands out particularly in past tense markers *d* > *dž* - *phendža* (*phend' a*) 'said', *t* > *č* - *dikča* (*dikh' a*) 'saw'.

Underlying velarization of *l* > *ɫ* appears here, as in the Romani dialects of southern Poland (cf. Rozwadowski 1936), with a pronunciation matching that of *ɫ* in contemporary Polish, i.e. [w]. The development may encompass underlying *l* in all positions except those preceding *i*, where *l* is preserved — *dudali* 'window', *marykli* 'cake', *bolipen* 'sky' (as well as in Polish words with *l*).

*kh* > *kš* in *kšit* 'butter' appears to be an isolated case, and it is unlikely that it reflects underlying Old Indic *kṣīr*-.

### 3.5. Summary of the Most Distinctive Phonological Features

Few phonological forms enable reconstruction of a particular underlying branch affinity in early Romani, beyond a very general grouping together of the Northern and some of the southern Balkan dialects. This in effect suggests the early separation of Vlax (Romanian-Transylvanian Romani), and its subsequent partial projection onto dialects now spoken in central Europe. Both the Northern and the

Central dialects could be viewed as western outposts of the southern Balkan branch, some of which have subsequently absorbed innovations spreading along the tighter geographical continuum of dialects in central Europe. The bulk of phonological developments attested in the dialect under consideration reflect specific geographical distributions of solutions to alternations and variations found in the language as a whole.

Perhaps the most accurate concise characterization of the position of the dialect may be obtained by locating it with respect to the following developments:

Unaffected by  $s = h$  (*si*, *sy* = *hi*, *hin* 'is', and elsewhere), found to various extents in the neighboring Sinti and the Central dialects, and deriving partly from a geographical spread in central Europe, and in part from underlying variation in the inherited Indic component.

Extremely consistent in avoiding initial *a* in thematic items, with either initial *ja* or  $\emptyset$ ; this coincides with third person pronouns in *j*-.

Shares  $as > a$  in the third singular past tense with other northeastern varieties, with palatalization of (underlying dental) past tense morphemes,  $d > d\check{z}/d'$ ,  $t > \check{c}/t'$ .

Shares the tendency toward vowel centralization to *y* with other Northern and especially northeastern dialects, while sharing  $l > t$  with other dialects having long established roots in Poland.

## 4. Morphology

### 4.1. Nominal Derivation

Northern derivational features include the full retention of *-pen/-ben* as the abstract nominal suffix. This thematic affix also appears in the plural and in non-nominative cases as well, where some dialects resort to athematic affixes from Greek: *čacipen* > *čačipnasa* 'truth > truthfully' (cf. Vlax *čačimasa*, based on athematic *-mos*), *xaben* > *xabena*, 'food > foods' (cf. Vlax *xabemata*), *hučipen* > *hučipnastyr* 'height > from above'. Adjectival endings may be used for secondary derivational purposes: *dud* 'light' > *dudali* 'window'.

Athematic endings are *-o* for masculine and *-a* for feminine: *mosto* 'bridge', *sveto* 'world', *konco* 'end', *banko* 'bench'; *klasa* 'class', *bilta* 'picture', *tyša* 'table', *planeta* 'planet'. The plural ending for both is *-y*: *sxody* 'steps', *dinosaury* 'dinosaurs', *bilty* 'pictures', *planety* 'planets'. As mentioned above, this coincides with neighboring Sinti to the west and Xaladitka to the east, while to the south we find masculine *-os/-is* as in *mostos* 'bridge' (Ackerley 1937; Kopernicki 1930).

### 4.2. Loss of Articles

As in the Xaladitka dialect described by Wentzel (1980: 134), underlying definite articles are used only with prepositions: *paši tyša* 'next to/at the table', *pašo*

*kher* 'next to/ beside the house'. Elsewhere, copying Polish, there is no usage of either definite or indefinite articles: *piravav dudali* 'I open the window', *dava sy čor?* 'is this a thief?'. This again contrasts with the southern Polish dialects, which retain articles.

#### 4.3. Case

On the whole and in comparison with other varieties, synthetic case endings that are retained are extremely productive, the dialect showing no ablative prepositions, while the total number of cases is reduced due to the disappearance of the locative.

The instrumental is used, as in other dialects, in a sociative/comitative meaning: *me rakirav mire matesa* 'I am speaking with my friend'. It is also extended to spatial and temporal relations: *pšegeja vudarensa* 'he passed through the door', *našty sovenys račasa* 'they couldn't sleep at night'.

As in other Northern dialects, long forms for the genitive prevail: *čhoneskry dud* 'moonlight', *dad mire mateskro* 'my friend's father'. Since ablative prepositions are missing, the tendency to paraphrase the genitive through prepositions is not observable here.

The accusative can be taken by all animates, including athematic nouns and indefinites: *dykhav myšos* 'I see the mouse', *dykhav biš grajen* 'I see twenty horses', *jov xtytəl mačhen* 'he catches fish', *me lav romes* 'I am getting married (=I take a man)'. Inanimate or abstract direct objects generally do not take the accusative: *denys teske but xabena* 'they gave him a lot of food (=many foods)', *me lubinav thud* 'I like milk'. However, the accusative functions consistently as the case of the possessor, with the topical prominence of the possessor taking precedence over animacy (cf. Matras 1997): as expected, *Jankos na sys kana zor* 'Janko had no more strength', but also *Saturnos sy but roty* 'Saturn has many rings'.

The dative is the case of the benefactive: *keret peske špery* 'he is creating tracks for himself', *me pekav mire dake marykla* 'I am baking cakes for my mother', *łove na čhenys tenge pro džipen* 'the money wasn't enough for them to live on'. It appears regularly only with a restricted number of verbs, including 'to give',<sup>8</sup> 'to show', and optionally with 'to say': *det čhavorenge te xal* 'she is giving the children to eat', *me sykhavav tuke varyso* 'I will show you something', *phendža lake kaj te jandel pani* 'he told her to fetch water', but *jov phendža ki peskry phen* 'he said to his sister'.

The ablative is dominant in expressions of source in the broadest sense. As in Xaladitka, and unlike most other dialects including neighboring Sinti and Central, its functions are not shared with an ablative preposition, which is missing: *vygeja kherestyr* 'he came out of the house', *vydžał vudarendyr* 'he is coming out

through the doors', *moxtenestyr až ki dudali* 'from the box to the window', *kerde sy barendyr* 'they are made of stones', *sys baxtalo frajdatyr* 'he was full of happiness (=merry from happiness)'.

An extraordinary feature of the dialect, not shared even with Xaladitka, is the disappearance of the synthetic locative. The expression of locative relations with the prepositions *ki* (non-incorporative) or *dro* (incorporative) resembles most other dialects of the language: *me džava ki vavir škola* 'I will go to another school', *sy dža čixo dro kher* 'it is so quiet in the house'. It is also not unusual in Romani for the locative to cease to function as a prepositional case with inanimates: *sys bešte paši tyša, glan čučē čare* 'they sat at the table, in front of empty plates'. However, with pronouns and animates in the dialect, it is the ablative that functions as a prepositional case, replacing the locative: *ki jamendyr* 'with us (=at our house etc.)', *jov džal ki mandyr* 'he is walking towards me', *pułapka pre myšendyr* 'a mousetrap (=trap for mice)'.

#### 4.4. Pronouns, Deictics, Indefinites

Noteworthy are some phonological developments reflected in the pronoun system. As noted above, initial *a-* is avoided in the dialect, and we find *jamen* 'we, us', *jamaro* 'our', alongside *menge* 'for us', *men* 'us'. Third person pronouns have, as in other Northern and Central dialects, *j-*forms: *jov* 'he', *joj* 'she', *jone* 'they'. The non-nominative paradigm for third person pronouns has initial *l-*: *teskro* 'his', *lake* 'for her'. The reflexive pronoun retains *-s* in *pes* throughout. With interrogatives, the series in *s-* is retained, unlike Sinti, where *h-* forms occur. The form for 'how much' is, as in Xaladitka, *kicy*.

Deictics are on the whole congruent with Xaladitka as described by Wentzel (1980), though it cannot be established at this stage whether a four-term system is in use: m. *dava*, *dova*, f. *daja*, *doja*, pl. *dała*, *doła*. The dialect is thus separated by an isogloss from both Sinti and the southern Polish (Central) dialects, which make use of the set of *k*-based deictics as well as those in *(a)d-*. A four-term system can be recognized based on the material at hand for place adverbs: *daj*, *adaj* 'here', *doj*, *odoj* 'there'. From the formal composition it would seem that the initial vowel prefix (*a-*, *o-*) here represents 'specificity' (cf. Matras 1998a), in line with the dialects of the southern Balkans.

A short form of the situational (extra-linguistic) deixis *da* appears to be used for continuous focusing, suggesting a development toward re-emergence of definite articles: *kaj si da bilta?* 'where is the well-known/that picture?'. *da* also forms the basis for comparative deictics as in *dasavo baro syr* 'as big as', *dakicy but syr* 'as many as', alongside *dža*, as in *dža but syr* 'as much as', or with the specific prefix *adža but?* 'that much?'.

Indefinites are based on *vary-*, shared with other Northern, but also with Central and Vlach dialects, while in Balkan varieties there is a strong tendency for indefinites to merge with those of the primary contact language: *varykon*, 'somebody', *varyso* 'something', *varysavo dyves* 'one day', etc. The presence of *vary-* here questions the etymology commonly suggested from Romanian *oare-*. There is no trace of *tšimone* 'something', which appears to be exclusive to the western sub-division within the Northern branch. The negative indefinite is *čhi* 'nothing'.

#### 4.5. Prepositions

On the whole the inventory is congruent with that of the Xaladitka dialect, showing some typical northeastern morpho-phonological developments, *maškyr* 'between', *pať* 'behind, for', *ki* 'at, to', and some general Northern forms, *dro* 'in', *pre* 'on, for', *pa* 'about', *paš* 'beside'. Noteworthy constructions are *až ki* 'up to, until' (Xaladitka and other dialects *ži ko*), *tať* 'below, under' (Xaladitka and elsewhere *tel*), *giť* 'around' (compare Xaladitka *angil* 'over, against'), and *głan* 'in front of, ahead of', found in Sinti, but not in Xaladitka.

The productivity of the synthetic ablative case is consistent with the lack, as in the Xaladitka dialect, of an ablative preposition comparable with *dral* found to the west in Sinti (alongside *von* from German), or *andar* and *katar* found to the south in the central dialects of southern Poland (Rozwadowski 1936) or Czechoslovakia.

#### 4.6. Adjectives

Northern formations are apparent in adjective derivation with *-itko* e.g. *panitko lvo* 'sea lion' < *pani* 'water', as well as with the use of genitive endings for derivation, e.g. *xolinakro* 'angry' < *xoli* 'anger'.

The dialect retains the more conservative synthetic comparative suffix: *butedyr* 'more' < *but* 'much', *tyknedyr* 'smaller' < *tykno* 'small', *fededyr* 'better', suppletive form for *mišto* or *lačho* 'good'. Contrasting with Xaladitka, where the superlative is formed by adding an adjectival ending to the comparative (*baredyro* 'biggest', cf. Wentzel 1980: 84), in the Polska Roma dialect the superlative shows the prefix *naj-*. This also appears the Carpathian dialect (Ackerley 1937) as well as in southern Balkan dialects, where it attaches to the simple adjective, while in Polska Roma it is added to the comparative forms: *najfededyr* 'best', *najbaredyr* 'biggest'.

As elsewhere, athematic adjectives take no gender, number, or case inflection: *optično vizja* 'an optical illusion'. With thematic items, the predicate adjective can take full case endings: *čhavoren na mukhasam korkoren* 'we shall not leave the children alone'. Adjectives show a weak declension in association with nominal case inflection on the adjoining noun. Unlike in the Xaladitka dialect, there is no

tendency to copy case inflection onto adjectives. Instead, gender and number are neutralized, and the adjective or possessive assumes a general oblique form: *jekh tykno vast* 'one small hand' > *zanta jekhe tykne vastestyr* 'sand (falling out) from one small hand' (ablative), *peskry romni* 'his wife' > *pučhel peskre romnjaty* 'he asks his wife' (ablative), *dava čhavo* 'this boy' > *me phendžom dave čhaveske* 'I said to this boy' (dative).

## 4.7. Verbs

### 4.7.1. Verb Derivation

A Northern feature is the productive derivation of verbs as pseudo-compounds with *-kirel*, derived from *kerel* 'to do': *sykhlakirel* 'to learn', *hačkirel* 'to understand'.

Passive derivation matches that of Common Romani. Synthetic passives are a derivation from the past tense form of the verbal root, which is palatalized (though in this dialect palatalization in the passive merges with the general palatalization of the past tense markers). In the present tense, personal endings are linked to the passive root by means of a back vowel *o* or *u*: *šunel* 'to hear', *šundžol* 'to be heard'. The past tense is formed with *-(i)ja*: *šundža* 'heard' > *šundžija* 'was heard'. As in other dialects, such formations are often active but non-agentive (*hačol*, *hačija* 'to shine'), and so it seems more appropriate to abandon the traditional notion of passive and speak instead, in the Romani context, of transitives, intransitives, and de-transitives (see also Minkov 1997). A curious item is the transitive meaning of the simple root in *resel* 'to achieve' (as in Xaladitka; but in other dialects *resel* 'arrive'), from which the verb 'to arrive' is formed by de-transitivization: *ryšol* 'to arrive'.

Medium passives seem to be especially productive and rather frequent in the dialect under consideration. Contact influence in this regard is probable, for reflexives are added to roots which in other dialects have intransitive meanings: *darav man* 'I am afraid', *šarav man* 'I rejoice'. Adding the reflexive pronoun is also the general procedure for 'genuine' de-transitivization: *sykhadža take syr pes dova kerel* 'he showed her how this is done', *vudara pirade pes* 'the doors opened'.

Athematic verbs are based on the root-extension *-in*: *malinel* 'to paint', *lubinel* 'to like', *moginel* 'to be able to', *trafinel* 'to find', *mišlinel* 'to think'. Athematic inchoatives add to the inchoative/de-transitive formation the extension *-uv*. Palatalization affects the root here as well: *frajdžuvet* 'to rejoice' > dialectal German *frajd* 'joy'.

### 4.7.2. Verb Inflection

The shortened present conjugation is used for the present tense, while the long forms in *-a* are used for the future tense: *kerav* 'I do', *kerava* 'I will do'. In this

the dialect differs from the pattern of retention of long forms for the present tense with an analytic expression for the future (Xaladitka, as well as southern Balkan dialects), or lack of a separate future altogether (Sinti). Instead we have a formation that figures in a geographical continuum encompassing this dialect, the southern Polish variety and other dialects of the Central branch, as well as Vlach. A rare conservatism is the retention of a consonantal ending in the first person plural: *mukhasam* 'we will leave', *džasam* 'we will go', likely to be a reflection of the Middle Indic future *-(i)ssāmo* (see Bubeník 1996: 104–105).

Noteworthy is the formation of a present tense based on the inflected participle for a number of verbs expressing state, as in *me bešto som* 'I (m.) am sitting', *jame sam sute* 'we are sleeping', *jov terdo sy* 'he is standing'. Alongside the retention of *-t* in some verbs, as in the aforementioned *bešto* and *sute*, we find a range of participle/past-tense affixes, though we do not find *-l*, which appears alongside *-n* for second and third persons plural in Xaladitka (*kamle/kamne* 'they loved'). The characteristic past tense formations are thus: *kerdža* pl. *kerde* 'did'; *rakhča* pl. *rakhte* 'found'; *phandija* pl. *phandyne* 'shut'; *lija* pl. *lyne* 'took'; *gija* pl. *gene* 'went'; *peja* pl. *pene* 'fell'; *xaja* pl. *xane* 'ate'.

The 2nd person singular ending on the copula and in past tenses is *-an*, contrasting with *-al* to the west (Sinti) and south (southern Poland; cf. Rozwadowski 1936, Kopernicki 1930). The "remoteness" suffix (which forms the imperfect from the present and the pluperfect/irrealis from the past, see Matras 1994: 69ff.) shows *-as > -ys*: *keretyš* 'he was doing', *mušindžamys* 'we would have had to'.

The modal verbs for 'can' and 'must' are (alongside impersonal *našty* 'cannot'), as in many other dialects, borrowed and so they show athematic inflection: *moginav* 'I can', *mušinav* 'I must'.

#### 4.7.3. Aktionsart Affixes

Like other Romani dialects in contact with Slavic languages, most notably Xaladitka and the Central dialects, but also some dialects in the Balkans, the variety under consideration borrows aspectual prefixes from its Slavic contact language. On the whole we are dealing with a modification of the verb root to render lexically specialized items, and so with the development of systematic Aktionsart distinctions: *but čiro pšegeja* 'much time passed' < *geja* 'went'; *syr rat zapeja...* 'when night fell...' < *peja* 'fell'; *odrakhte drom kaj zaligirdža ten ki tengro kher* 'they found the road that brought them home' < *rakhte* 'they found', *girdža* 'led'; *čirikle vixane o maro* 'the birds ate up the bread' < *xane* 'ate'; *jone podšunenys paš vudar* 'they were listening at the door' < *šunenys* 'they heard'.

Only in some cases does the aspectual modification bear a direct spatial semantic implication derived from the underlying meaning of the affixes: *dodav*

*cukro* 'I add sugar' < *dav* 'I give'; *odlel* 'to answer the telephone' < *lel* 'to take'; *vygeja kherestyr* 'he came out of the house' < *geja* 'went'; *bara sys rosčurdyne* 'stones were scattered' < *čurdyne* 'thrown'. Nevertheless, the potential attachment of any Slavic affix onto any Romani verb despite the availability of other means to express spatial or punctual relations, or indeed the lack of a universal necessity to do so, must be assessed as a wholesale incorporation of the system of affixed aspectual modification into the language. The correct term to designate this phenomenon is in my opinion "fusion" (see Matras 1998b). It implies that the two systems have become inseparable to speakers. For the purpose of processing and verbalizing the directional and durational conceptualizations associated with actions and events, both systems share the resources provided by the contact language.

The presence of a productive set of Slavic aspectual affixes constitutes a clearly identifiable isogloss within the dialects of Romani. Less clear however is the extent of areal diffusion of the plain concept of affix modification of the verbal root to indicate Aktionsart-relevant properties such as punctuality, containment, or directionality. Varieties of Sinti adopt a whole range of so-called verbal particles from German, mostly those expressing concrete spatial relations—*me džau hin* 'I am going [directional]'—, while calquing others—*kras o vuder pre* 'he opened the door' [lit. 'made it up' < German *machte auf*]. Such calquing of structural elements expressing Aktionsart can also be observed among Vlax speakers who are first generation immigrants in Germany: *asal tut avri* 'he is laughing at you' < German *lacht dich aus*. It appears that conceptualizations in terms of specific spatial metaphors of event duration, punctuality, or outward projection of an action are easily susceptible to transfer and replication in language contact situations. The progression of this transfer, from calquing through replication of isolated spatial morphemes (as in the case of Sinti *hin*) and on to, finally, the wholesale adoption (fusion) of the class of verbal modifiers from Slavic has yet to be explored.

#### 4.8. Negation

The Polska Roma dialect follows the pattern of the non-Vlax branches of Romani, with *na* as the principal verb negator (*na pačav* 'I don't believe') and *nani* as the negated third person copula. The imperative is negated by *ma* (*ma rov!* 'don't cry!'). An impersonal negated modal *našty* 'cannot' appears alongside *na moginel* 'not to be able to'. The negative indefinite is *čhi*, which is not capable of negating predications and so it combines with other forms of negation: *nani men čhi* 'we have nothing', *čhi nani te šundžot* 'nothing is to be heard', *dava lenge čhi na dija* 'this did not render them anything'.



#### 4.9. Adverbs

Lexical adverbs drawing on Common Romani *-es* may describe actions (*tokhes* 'softly', *čoraxanes* 'secretly'), but also states, where they are sometimes preferred over adjectives: *banges* 'crooked', *hučes* 'above'. Location adverbs correspond to those of other Romani dialects: *avri* 'outside', *maškyral* 'in-between', etc.; repetition is expressed by *pate* 'again'. A particular feature of the Polska Roma dialect is forms in *na-*: *najevkar* 'often' < *na jekh var* lit. 'not one time', *nadaryg* 'suddenly'. Adverbial focus particles, as is usually the case in Romani, are borrowed from the contact language: *tylko* 'only', *ceto* 'all', *tyš* 'too/also', *za* 'too (much)'.

### 5. Syntax

A common view expressed especially in older descriptions is that Romani dialects show no more than a minimum of independent syntax, drawing on their respective contact language for syntactic relations. This is at least in part inspired by the fact that coordinating conjunctions (and other discourse markers) and partly word order rules are likely to undergo convergence with a dominant contact language after several generations. While the view that Romani has no independent syntax is strongly challenged in Matras (1994), where a series of syntactic formations are argued to be the outcome of Balkan areal convergence in early Romani, the diversity of Romani discourse markers is explained in Matras (1998b) in universal terms. The present section outlines some outstanding features of the dialect while pointing out items that could serve as coordinates for a comparative framework of investigation of syntactic features in Romani varieties.

#### 5.1. Coordination

Coordinating conjunctions, along with other discourse markers, are in many dialects of Romani candidates for fusion with the structures of the contact language. Their wholesale replacement through Polish items in the dialect under consideration constitutes nevertheless a rather extreme case; there is no trace even of the Romani conjunction *ta* 'and', which has survived in more conservative forms of other Northern varieties. With the adoption of Polish items the dialect also adopts the Polish coordinative configuration, for example a three term additive-contrastive continuum with Polish *i* 'and' (addition), *a* 'and however' (semantic opposition), and *ale* 'but' (denial of expectation). Other items include *čy* 'whether, or, either', *to* 'then, so then', *bo* 'because'.

## 5.2. Adverbial Subordination, Embeddings, and Relative Clauses

As in other dialects of Romani, subordination is introduced by conjunctions/relativizers derived from interrogatives. Noteworthy is the absence of a temporal conjunction *kana* 'when', and the use instead of *so* 'what' for general simultaneity: *so javava phuredyr* 'when I become older', *me frajdžuvav so dad pale ryšoť* 'I am happy when father comes back'. Specific (immediate) simultaneity is expressed, as in other dialects, with *syr*: *syr rat zapeja* 'when (=just as) night fell', *dykhav myšos syr pro moxten phireť* 'I see the mouse running on the box'. Conjunctions of the same type introduce embeddings as in *dikhel syr thulo jov sy* 'he sees how fat he is', *najevkar na sys ten so te xat* 'they often had nothing to eat', *džinav so kana kerava* 'I know what to do now'. The latter examples also serve as an illustration of the prominence of finiteness in the language.

Relative clauses are introduced as in other dialects with *savo*, *kaj*, or *so*: *love save jov zaxitetyś* 'the money which he was earning', *bara kaj pašo kher sys* 'the stones that were next to the house', *saro so pesa lija kherestyr sys maro* 'all he took with him from home was some bread'. The conditional conjunction is as elsewhere *te*.

## 5.4. Complementation

In regard to a comparative approach to the syntax of Romani dialects, complementation is, alongside word order, perhaps the most exciting domain. We are concerned here essentially with two features: 1) the rules on verb agreement with coreferent subjects in modal constructions, and 2) the distribution of complementary conjunctions along the factuality (indicativeness) continuum.

### 5.4.1. Verb Agreement with Coreferent Subjects

Let us first establish that Romani is on the whole an overwhelmingly finite language. The tendency in some Romani dialects to adopt solutions other than finite agreement on the verb in modal complements has been referred to by Boretzky (1996) as the "new infinitive." The term captures an historical dimension to be seen in connection with a partial de-balkanization of Romani dialects outside the Balkans, where the famous reduction or loss of the infinitive occurs (cf. Joseph 1983). The underlying assumption, supported by the actual data and structural features of the so-called "new infinitive," is that even those Northern and Central dialects in which a "new infinitive" occurs once lacked an infinitive as a result of the effects of the process of syntactic convergence among the Balkan languages on common or early Romani.

In actual fact, even de-balkanized dialects of Romani, such as the one under consideration here, do not exhibit true infinitive forms. Rather, the term chosen by Boretzky (1996) pertains to the generalization of a form from within the finite paradigm. Moreover, this form is introduced by the non-factual complementizer *te* thus granting the complement full clause status, rather than allowing it a nominal-like treatment. It therefore seems preferable simply to speak of agreement versus non-agreement of the complement verb with the co-referent subject of the main clause.

In the dialect under consideration, non-agreement is characterized by a third person present-tense or “neutralized” inflection on the subordinated verb: *so som kana te keret?* ‘what am I to do now?’, *so sam te det čhavoreng te xa?* ‘what are we to give the children to eat?’ A condition for non-agreement however is, unsurprisingly, the presence of a formal marking of the subject antecedent at least in the inflection of the main verb, as in the above examples, if not through the presence of a subject NP. Thus non-agreement appears in *me kamav tes te urakhet* ‘I want to free him’, but not in *našty tes te urakhav* ‘I cannot free him’, where the main verb is impersonal and no antecedent is present.

#### 5.4.2. The Factuality Continuum

The distribution of complementizers is characterized firstly by two opposite poles on a continuum of factuality and event independence: *te* is the complementizer that introduces modal complements (or more generally, non-factual predications, including optatives and first person imperatives). These show tight semantic integration of one single event—*najevkar na sys ten so te xa?* ‘sometimes they didn’t have anything to eat’—and in the case of personal modals, a subject coreferential with that of the main clause: *jame moginas dava te zumavet* ‘we can try this’, *ki Marso mušinet duj berš raketa te tradet* ‘a rocket has to travel two years (to get) to Mars’, *lubinav man te pieščinet* ‘I like to cuddle’, *kamas te kinet pułapka pre myšendyr* ‘we want to buy a mousetrap’. Characteristic for complements introduced by *te* is the weak independent truth value of the predication they represent and so their semantic dependency on the main verb. In this regard *te* can be seen as introducing subjunctive clauses, even if, as in this dialect, no subjunctive conjugation actually appears on the verb.

At the other end of the continuum, *kaj* is used with factual relations, in particular complements of epistemic predications: *dava sy tylko paramiši kaj pro Marso manuša sy* ‘it is only fairy-tales that there are people on Mars’, *phendža kaj na džinet* ‘she said she didn’t know’, *dikhča kaj takro pšať čhija bara* ‘she saw that her brother threw stones’, *džindža kaj čovahani kamet ta te xa?* ‘she thought that the witch wanted to eat her’. Factual complements represent events that are potentially

independent of the main verb, hence their attachment to verbs of perception and expression, as opposed to volition and manipulation, which take non-factual complements. Thus *kaj* can be regarded as a marker of indicative relations.

On the whole this corresponds to the pattern shared by the Balkan languages, hence one may continue to regard the Polska Roma dialect as Balkan in regard to its complementation typology.<sup>10</sup> The difficulty lies in defining the distribution of complementizers in the area in between the two extremes on the continuum. The relevant constructions where variation can be assumed to occur are those where event integration is less straightforward than in modal constructions, and where on the other hand event independence and factuality are less obvious than with epistemic verbs. This pertains to two types of constructions: purpose clauses and manipulative clauses. In these domains there is variation between plain modality in *te* and what may be called “enforced modality,” expressed by a combination of two conjunctions: *kaj te*.

Plain modality in *te* appears in “permissive” manipulation: *na defys take te xat* ‘he gave her nothing to eat’. The subjects of the two parts of the complex construction differ and so events are potentially independent of one another. On the other hand, relatively little use of manipulative force is necessary in order to exercise sufficient control that would allow the second subject to carry out the action indicated in the embedded clause.

With true or “active” manipulation, the subjects differ and considerable manipulative authority is needed on the part of the first subject in order to obtain the intended results; this is where *kaj te* is inserted: *phenetys ki lestyr kaj te sykhavet take peskro vast* ‘she told her to show her her arm’. However, in chainings of embedded verbs in such “active” manipulative constructions only *te* is needed: *phendža take kaj te jandel pani, i te keret jag andro bov* ‘she told her to fetch water, and to light a fire in the stove’. Thus, once the nature of the manipulative relation between the actions described in both parts of the complex construction has been established, only the strict modality component needs to be repeated in the second, conjoined predication of the embedded complement. It follows that the stronger the manipulative force involved (authority, persuasion), the more the action portrayed in the embedded clause is perceived as potentially independent.

The second construction type which is potentially ambivalent is the purpose clause. Here, the same subject can be involved in both parts of the construction, but the actions are typically independent. Once again the type of complementation depends on the degree of integration. I distinguish between “direct intent” and “reflective intent.” With “direct intent,” the embedded predication is a natural and expected outcome of the main predication: *taša džasam sare dro veš kaša te čhinet* ‘tomorrow we will all go to the forest to cut wood’. With “reflective intent,” there is no natural connection between the background event and its target action, and

more force is required in order to evoke the desired outcome: *varyso kerasam kaj men dalestyr te urakhas* 'we will do something to free ourselves from this'. For example, there may be circumstantial constraints on the subject which may affect his ability to control the outcome of the embedded predication: *Jankos na sys kana zor kaj pale te ponaskendet bara* 'Janko no longer had the strength to gather stones'.

Finally there is a mixed construction type, the manipulative purpose clause, or "indirect manipulation" (=final construction), where the subject of the main clause exercises its manipulative force on itself, as in a purpose clause, but does so in order to evoke an action on the part of a second subject: *phendža tokhes ki peskry phen, kaj dada te na šunen* 'he said gently to his sister, so that the parents will not hear'. This type combines properties of both purpose and manipulative clauses; the employment of enforced modality in *kaj te* can be explained by the need to exercise active manipulative control.

Note from the point of view of word order that the complex complementizer *kaj te* may be split. This can be expected from the general rule in other Romani dialects (see Matras 1995) according to which the non-factual complementizer always triggers connective (i.e. verb-initial) word order. When elements are fronted, that is, inserted into pre-verbal position, they necessarily appear before *te*, which is inseparable from the verb. At the same time *kaj* remains fixed to the position marking the begin of the subordinated predication.

In Matras (1994: 234-236), a corresponding structure in a Vlax dialect, *ke te*, was defined as an attempt by the speaker to provide a logical and presuppositional justification for the relation between background action (main clause) and intended outcome (embedded purpose clause). In other words, the "enforced" non-factual embedding with *ke te* (or *kaj te* in the dialect under consideration here) indicates that the background action described in the main clause cannot unconditionally be viewed as a sufficient prerequisite for the successful realization of the target action described in the embedded complement. In accordance with the functional-pragmatic approach taken in that study, variation among *te* and *ke te* was interpreted as a function of the speaker's assessment of the degree of shared presuppositional knowledge concerning the particular logical chain [background action > intended outcome]. The more obvious the chain is to both speaker and hearer, i.e. the more presuppositional it is, the more likely it is for the plain subordinator *te* to be employed. The use of *ke te* in the Kelderaš/Lovari variety is confined largely to argumentation structures, and so it is more restricted than the use of *kaj te* in the Polska Roma dialect. It is nevertheless apparent that the borderline between the two structures is open to variation based on speaker's judgements, for which pragmatic considerations of the kind sketched above play a crucial role.

### 5.5. Word Order

In assertive main clauses, the verb follows thematic<sup>11</sup> elements and is followed by rhematic ones (SVO): *čhavoro džat ki škola* 'the boy goes to school', *me dikhav biš grajen* 'I see twenty horses'. There is a tendency in both simple and complex constructions to front object and reflexive pronouns to preverbal position: *me man šarav* 'I am happy'. This may apply to indirect pronouns and to indefinites as well: *dava tenge čhi na dija* 'this didn't give them anything'. There is however no strict formal rule to this effect, and constructions such as the following are also encountered frequently: *jov pučhel pes doxtorostyr* 'he is asking the doctor for advice'.

Verb-initial order appears in presentative constructions: *bešto sys čororo manuš* 'there lived a poor man'. The vacant thematic position can be interpreted pragmatically as a start-from-void where speaker and hearer share no previous information about relevant thematic entities, hence the employment of this device to initiate pre-planned narratives. Verb-subject order is also found in chained predications where it indicates some form of consequentiality (= "connective" word order; cf. Matras 1995): *vudara pirade pes i vygeja tengro dad* 'the doors opened and out came their father' (cf. also discussion in Holzinger 1993).

Verb-subject order is common in interrogative clauses: *rakhela myšo da drom?* 'will the mouse find the way?', *sy da kreski bange?* 'are these circles bent?'; and when an interrogative appears in first position: *so keren sare čhavore?* 'what are all the children doing?', *kaj sy da bilta?* 'where is that picture?'. But once more there is no strict formal rule to this effect, and interrogative clauses can just as well contain thematic subject-verb order: *so dava sy?* 'what is this?', *syr bare jone sy?* 'how big are they?'. Object pronouns and nominal objects may be fronted: *so pes adaj keret?* 'what is going on (=lit. is being done) here?', *kon miro kher phagiret?* 'who is breaking my house?'.

A distinctive feature is the tendency to place the verb in the final position of subordinated clauses, fronting objects of various kinds: *mire kana moginav me tat pani ten te phandel* 'I can shut my eyes below the water', *dykhav myšos syr pro moxten phiret* 'I see the mouse walking on the box', *na dikhetys mišto so jov take sykhavel* 'she couldn't see well what he was showing her', *syr ten dikhča, sys baxtało frajdatyr* 'when he saw them, he was full of joy'.

### 6. Conclusion

From the above remarks a number of salient features may be derived, allowing a brief characterization of the position of the Polska Roma dialect.

In phonology it is rather extreme in not allowing initial *a-* in thematic roots, and in not allowing *h* as a substitute for *s* in morphological paradigms. In morphophonology it shows third person pronouns in *j-*, it contracts the past tense third singular suffix from *-as* > *-a*, and it does not show a past-tense resultative (perfect, or “aorist”) infix in *-l-* (retaining instead *t* > *č*); in morphology it shows a loss or at least extreme reduction of definite articles, it replaces the locative with the ablative case, it does not participate in *-an* > *-al* for the second singular of copula and past tense conjugations found in neighboring Sinti and the Central branch, it uses “short” forms for the present conjugation and “long” forms for the future, its use of participles for active tenses is restricted to some present-tense formations, and it adopts the set of Slavic aspectual prefixes to indicate Aktionsart; in morphosyntax it shows no ablative preposition, it abolishes verb-agreement with anteceded coreferential subjects in modal complements, it tends to front pronominal direct objects and to place the verb in final position in subordinate clauses, and it relies on a combination of indicative+subjunctive complementizers in complex constructions where manipulative control is semantically ambivalent.

How can this set of features be interpreted in the broader context of Romani dialectology? From our above discussion it appears that only few traits can be classified as “Northern” and found to correspond both to the North Russian (Xaladitka) and Baltic dialects to the east, and to Sinti-Manuš to the west (both “Northern” affiliations). “Northern” features are restricted to a relatively modest inventory of lexical items, a number of phonological developments, and perhaps also the structure of athematic morphology. In respect to the majority of other developments sketched in the above outline, the dialect can be shown to form part of a geographical continuum. We might therefore imagine an underlying Northern group to have shared some phonological innovations (or conservatisms), some lexical particularities, as well as a distinct pattern of athematic incorporation, which in turn suggests dialect division within early Romani.<sup>12</sup> On the whole however these features are less significant for subsequent development. More important is the geographical spread of grammatical innovations, including morphophonological and morphosyntactic features. To this one must add convergence (in the broadest sense of the term) with coterritorial languages, or put more plainly, contact influence.

To the features figuring in our discussion with apparent areal diffusion within Romani we can reckon, among others, the following: 1) The *s/h* isoglosses (with at least two distinct centers of projection, one affecting Sinti-Manuš and the Central varieties, and one in southeastern Yugoslavia); 2) the diffusion of third person pronouns in *j-*, 3) the specialization of “long” conjugations for future tense, 4) the loss of verb-agreement with coreferential subjects in modal complements, 5) the forms *-an* versus *-al* for the second person singular of copula and past tense

conjugations, and 6) the retreat of past participles from the formation of third person active past tenses with intransitive or unaccusative verbs and the generalization instead of inflected forms. The position of the semantic continuum between *te* and *kaj* subordinations, the degree to which there is overlap among dialects in this domain and the question whether areal diffusion is involved remain to be established.

Alongside these internal developments with an areal spread, there are contact-related phenomena which in effect show areal correlations, most notably the emergence of Aktionsart differentiation through the wholesale adoption of the Slavic system of prefixes. As noted above, the adoption of semi-bound verbal particles from German in Sinti-Manuš could constitute a development worth examination within the same context of contact-related areal diffusion. A further, though more restricted areal phenomenon triggered by convergent patterns is the loss of articles, which forms a Polish - North Russian - Baltic continuum. Strictly contact-related phenomena represented in the dialect under consideration are, disregarding lexical borrowing, the system of coordination and discourse marking (or "utterance modifiers," see Matras 1998b), formalization of word order alternations and especially changes in the position of thematically "weak" elements such as object pronouns and adverbs, the use of borrowed modals, and phonological changes.

Finally there is a set of structural properties that represent dialect-particular innovations from what might go back to an underlying variation within earlier stages of historical Romani. Among its numerous phonological manifestations we find a variety of reflections of initial *a-/Ø-* alternation (*arakh-/rakh-* etc.), with solutions particular to individual words in individual dialects. Variation in the domain of deictic expressions might be considered in this connection, with individual dialects choosing from an inventory of inherited forms. Dialect-particular leveling of inherited variation need not however be considered in isolation. Especially in relation to deictics there is evidence that dialects have borrowed from one another (cf. Boretzky 1995), so that internal regulation of variation will have interacted with earlier branch division, as well as, at a later stage, with interdialectal contacts. Moreover, in relation to causativization it has recently been argued that external contact (i.e. with coterritorial languages) may reinforce the retention and productivity of inherited morphology (Hübschmannová and Bubeník 1997).

A further phenomenon worth considering in this context is the very gradual trend toward analytic case formation at the expense of synthetic case marking. Here too, dialects adopt different individual solutions as far as retention of synthetic (Layer II) case markers is concerned (cf. Matras 1997). Our present discussion offers some clues with respect to the origins of the process. Consider those prepositions that are in direct competition with synthetic case markers, that is, those that add minimal semantic specification to the local relation expressed by the case



suffix. All Romani dialects have developed two prepositions expressing a locative relation, an incorporative *and-/andr-/dr-* (containing predication) and a non-incorporative *k-* (non-containing predication). Conversely, no Romani dialects have developed either dative prepositions (though locatives, serving as directionals, occasionally substitute for datives) or instrumental/comitative prepositions. Variation occurs in the domain that appears, as far as the stability of synthetic markers is affected, in an intermediate hierarchical position (see Matras 1997): ablatives. The plane of Romani speech communities that comprises the southern Balkan, Vlax, and Central branches shows incorporative *andal/andral* alongside non-incorporative *(ka)tar*. Sinti-Manuš only has an indigenous incorporative *dral*, and uses German *von* for non-incorporative. The northeastern varieties, Polska Roma and Xaladitka, have no ablative preposition at all. This picture suggests a split among Romani dialects during a stage at which only locative prepositions had become fully grammaticalized, reinforcing once again the impression that the emergence of prepositions is connected to the process of syntactic Balkanization during the constitutive phase of “early European Romani” (cf. Matras 1994).

Much like the comparative description of phonological and morphological processes in Romani dialects, a much lacking morphosyntactic typology of Romani dialects must therefore draw on the various layers of inherited traits, features that evolved through geographical diffusion and convergent development among Romani dialects, and features that are a result of convergence with non-Romani coteritorial languages. From our present discussion, the following features may be identified as a partial agenda for such a comparative investigation of morphosyntactic developments in Romani: 1) the patterns of athematic derivational morphology, 2) the status of pseudo-compounds in verbal derivation, and of compound (genitive) nominalizations for nominal derivation, 3) overt marking of definiteness and indefiniteness, 4) the distribution of transitivizing and de-transitivizing options among verbal roots, 5) consistency in the tense formation of the verbal system, and in particular the status of past participles in the formation of active tenses, 6) grammaticalized expression of Aktionsart, 7) the distribution of synthetic case and prepositions, 8) verb agreement in modal complements with coreferent subjects, 9) expression of ambivalent control over embeddings (factuality continuum), 10) formalization of word order rules, and 11) renewal through borrowing of the inventory and distribution of conjunctions.

In conclusion, it can be said that the material considered here shows some evidence for a division within early Romani, as well as for carrying-over of variation from early Romani with subsequent different developments in individual dialects. In addition, it is safe to assume cross-dialectal interference at earlier stages. All these considerations will account for features of the Polska Roma dialect that are shared on a wholesale basis with the Northern branch, and more remotely with southern

Balkan varieties. A series of other developments, however, might be related to the areal diffusion of innovations, often encompassing the Vlax and Central branches along with the Polska Roma dialect, or alternatively shared only with the Central branch. This places the Polska Roma dialect at the westernmost edge of a plane of linguistic innovations. Furthermore, it suggests that despite the absence of settlement patterns covering coherent territories, Romani speech communities might indeed be regarded as a geographical continuum that allows for the spread of structural innovations.

## Notes

*Acknowledgment.* I owe special thanks to Peter Bakker for helpful remarks on an earlier version of this paper.

<sup>1</sup> One might add scarce evidence that vocabulary may have occasionally entered some Northern dialects from Rotwelsch-like jargons. One such item is *Štaxelengero* ‘hedgehog’ in the dialect of the German Sinti, from Rotwelsch *Stacheling*.

<sup>2</sup> The Polish Vlax population, mainly Lovara, use *rumungri* to refer to the *polska roma*, which perhaps suggests a perception of the two groups as closely related linguistically.

<sup>3</sup> Papers presented by N. Boretzky and by M. Hübschmannová et al. at a November 1997 meeting in Vienna on “Roma und Öffentlichkeit.”

<sup>4</sup> I am grateful to Andrzej Wisniewski for his assistance in obtaining a significant part of the data.

<sup>5</sup> “Athematic” morphology is the cover term for derivational affixes added to the lexical component acquired during or after the Byzantine period, while “thematic” is reserved for the Asian lexicon.

<sup>6</sup> In some dialects, however, a phonemic distinction between trilled and uvular /r/ is retained.

<sup>7</sup> Alternations such as *an-ljan* ‘to bring’, *av-ljav* ‘to come’, *aver/javer* ‘other’ are found in the Carpathian dialect (Kopernicki 1930; also Ackerley 1937).

<sup>8</sup> In Vlax dialects “to give” frequently governs the accusative.

<sup>9</sup> Compare *maj* in Vlax, from Romanian *mai* ‘more’. As with the indefinite *vary-*, it is not quite clear whether we have an underlying Romanian form here. If this were the case, then one would need to assume a very early layer of Romanian grammatical loans that preceded the secession of the Vlax dialects. There is little evidence to support such an assumption.

<sup>10</sup> This Balkan feature extends to the Balkan fringe language Serbo-Croatian, but a separation of complementizers of this type is also found in Czech.

<sup>11</sup> “Thematic” here refers of course to the distinction between theme and rheme in functional sentence perspective, and not to the term used specifically to refer to the morphological split between inherited and borrowed inflection in Romani.

<sup>12</sup>The significance of athematic morphology should not be underestimated. If Bakker (1997) and Igla (1997) are correct in deriving athematic morphology from Greek derivational patterns, then the distinct athematic patterns found in different Romani dialects might reflect a division within Romani during the phase in which contact with Greek was still productive for Common Romani, and so before major diversification through migration away from the Greek-speaking area.

### References cited

- Ackerley, Frederick G. 1937. The Polish Romani Vocabulary of Izydor Kopernicki. *Journal of the Gypsy Lore Society* 3, 16: 33-83.
- Bakker, Peter. 1995. Notes on the Genesis of Caló and other Iberian Para-Romani Varieties. *In* *Romani in Contact: The History, Structure and Sociology of a Language*. Yaron Matras, ed. Pp. 125–150. Amsterdam: Benjamins.
- . 1997. Athematic Morphology in Romani: The Borrowing of a Borrowing Pattern. *In* *The Typology and Dialectology of Romani*. Yaron Matras, et al., eds. Pp. 61–93. Amsterdam: Benjamins.
- . 1999. The Northern Branch of Romani: Mixed and Non-mixed Varieties. *In* *Das österreichische Romani im europäischen Kontext*. Dieter Halwachs and Florian Menz, eds. Wien: Böhlau. In press.
- Bakker, Peter, and Hein van der Voort. 1991. Para-Romani Languages: An Overview and Some Speculations on Their Genesis. *In* *In the Margin of Romani: Gypsy Languages in Contact*. Peter Bakker and Marcel Cortiade, eds. Pp. 16–44. Amsterdam: Institute for General Linguistics.
- Bakker, Peter, and Yaron Matras. 1997. Introduction. *In* *The Typology and Dialectology of Romani*. Yaron Matras, et al., eds. Pp. vii–xxx. Amsterdam: Benjamins.
- Bartos, Adam. 1994. The Social and Political Status of the Roma in Poland. *Roma* 40: 16–23.
- Boretzky, Norbert. 1995. Interdialectal Interference in Romani. *In* *Romani in Contact: The History, Structure and Sociology of a Language*. Yaron Matras, ed. Pp. 69–94. Amsterdam: Benjamins.
- . 1996. The “New” Infinitive in Romani. *Journal of the Gypsy Lore Society* 5, 6: 1–51.
- Boretzky, Norbert, and Birgit Igla. 1991. Morphologische Entlehnung in den Romani-Dialekten. Arbeitspapiere des Projektes “Prinzipien des

- Sprachwandels, "Nr. 4. Essen: Universität Essen, Fachbereich Sprach- und Literaturwissenschaften.
- . 1994. Romani Mixed Dialects. *In* *Mixed Languages: Fifteen Case Studies in Language Intertwining*. Peter Bakker and Maarten Mous, eds. Pp. 35–68. *Studies in Language and Language Use*, 13. Amsterdam: IFOTT.
- Bubeník, Vít. 1996. The Structure and Development of the Middle Indo-Aryan Dialects. Delhi: Motilal Banarsidass.
- Dowojno-Sylwestrowicz, Mieczysław. 1889. The Lithuanian Gypsies and Their Language. *Journal of the Gypsy Lore Society* 1, 1: 251–258.
- Ficowski, Jerzy. 1950. The Polish Gypsies of To-day. *Journal of the Gypsy Lore Society* 3, 29, 92–102.
- Finck, Franz Nikolaus. 1903. *Lehrbuch des Dialekts der deutschen Zigeuner*. Marburg: Elwert.
- Günther, L. 1915. Vorbemerkung (Preface). *In* *Die jenische Sprache*. Engelbert Wittich. *Archiv für Kriminal-Anthropologie und Kriminalistik* 63: 1–22.
- Holzinger, Daniel. 1993. Das Rómanes. *Grammatik und Diskursanalyse der Sprache der Sinte*. Innsbruck: Institut für Sprachwissenschaft.
- Hübschmannová, Milena, Hana Šebková and A. Žigová. 1991. *Kapesní slovník romsko český a česko romský*. Praha: Státní pedagogické nakladatelství.
- Hübschmannová, Milena and Vít Bubeník. 1997. Causatives in Slovak and Hungarian Romani. *In* *The Typology and Dialectology of Romani*. Yaron Matras, et al., eds. Pp. 133–145. Amsterdam: Benjamins.
- Igla, Birgit. 1997. Romani. *In* *Kontaktlinguistik*, Vol. 2. Hans Goebl, Peter H. Nelde and Zdenek Star, eds. Pp. 1961–1971. Berlin: Mouton de Gruyter.
- Joseph, Brian. 1983. *The Rise and Fall of the Balkan Infinitive*. Cambridge: Cambridge University Press.
- Kaufman, Terrence. 1979. *Review of Multilingual Romani Dictionary*, by W. R. Rishi. *In* *Romani Sociolinguistics*. Hancock, Ian, ed.. *International Journal of the Sociology of Languages* 19: 131–144.
- Kopernicki, Izidor. 1930. *Textes tsiganes*. *Mémoires de la Commission Orientaliste*, 7. Kraków: Polish Academy of Sciences.
- Ladefoged, Jakob. 1998. Romani elements in Non-standard Scandinavian Varieties. *In* *The Romani Element in Non-standard Speech*. Yaron Matras, ed. Pp. 133–163. Wiesbaden: Harrassowitz.
- Lerch, Hans-Günter. 1976. *Das Manische in Gießen. Die Geheimsprache einer gesellschaftlichen Randgruppe, ihre Geschichte und ihre soziologischen Hintergründe*. Gießen: Anabas-Verlag.
- Liebich, Richard. 1863. *Die Zigeuner in ihrem Wesen und ihrer Sprache*. Leipzig: Brockhaus.

- Matras, Yaron. 1994. Untersuchungen zu Grammatik und Diskurs des Romanes. Dialekt der Kelderaša/Lovara. Wiesbaden: Harrassowitz.
- . 1995. Connective Word Order in Romani. *Sprachtypologie und Universalienforschung* 48(1/2): 189–203.
- . 1996. Prozedurale Fusion. Grammatische Interferenzschichten im Romanes. *Sprachtypologie und Universalienforschung* 49(1): 60–79.
- . 1997. The Typology of Case Relations and Case Layer Distribution in Romani. *In* The Typology and Dialectology of Romani. Yaron Matras, et al., eds. Pp. 61–93. Amsterdam: Benjamins.
- . 1998a. Deixis and Deictic Oppositions in Discourse: Evidence from Romani. *Journal of Pragmatics* 29: 393–428.
- . 1998b. Utterance Modifiers and Universals of Grammatical Borrowing. *Linguistics* 36(2): 281–331.
- (ed.). 1998c. The Romani Element in Non-standard Speech. Wiesbaden: Harrassowitz.
- Miklosich, Franz. 1872–1880. Über die Mundarten und Wanderungen der Zigeuner Europas. Vienna: Karl Gerold's Sohn.
- Minkov, Michael. 1997. A Concise Grammar of West Bulgarian Romani. *Journal of the Gypsy Lore Society* 5, 7 55–95.
- Rozwadowski, Jan. 1936. Wörterbuch des Zigeunerdialekts von Zakopane. Krakow: Nakladem Polskiej Akademji Umiejetnosci.
- Sampson, John. 1926. The Dialect of the Gypsies of Wales. Oxford: Clarendon.
- Sergieviski, M.N. 1931. Cyganski Jazyk. Moskva: Central'noe Izdatel'stvo Narodov SSSR.
- Smart, Bath Charles, and Henry Thomas Crofton. 1875. The Dialect of the English Gypsies. London: Asher.
- Wagner, Max L. 1937. Stray Notes on Spanish Romani: Chapter 2, Cryptolalic Formations in other Romani Dialects. *Journal of the Gypsy Lore Society* 3, 16 27–32.
- Wentzel, Tatjana. 1980. Die Zigeunersprache. (Nordrussischer Dialekt). Leipzig: Enzyklopädie.