

Giulia MELI (Milano)

## REMOTENESS MARKERS IN KALAJDŽI ROMANI AS SPOKEN IN MONTANA (BULGARIA)

**Abstract:** Romani imperfect and pluperfect are built by the agglutination of the same morpheme to the inflected forms of the present and the perfect, respectively. This morpheme, labelled as “remoteness marker” (Matras 2001: 35) by the literature on Romani varieties, conveys a temporal value of distance towards a determined point of reference excluding at the same time any overlapping with the moment of speech, and thus its meaning approaches the “temporal discontinuity” highlighted by Plungian and Van der Auwera (2006). The remoteness marker is quite homogeneous in Romani varieties and the main recorded forms in the different dialects (-*as*/*-a*/*-e*/*-s*/*-ys*/*-ahi*, cf. Matras 2002: 152) allow to reconstruct a single Proto-Romani form \**asi* (cf. Bloch 1932, Bubeník 1995) or \**sasi* (Scala 2020), both going back to the Old Indo-Aryan *as-* ‘to be’, maybe through the third person Middle Indo-Aryan form *āsi* or *āsī* ‘he/she/it was’. Nevertheless, some dialects show a greater complexity and a certain level of internal variation, and suggest that the general uniformity displayed by Romani varieties may have been preceded by a more composite situation. In particular, the paper analyses the remoteness markers of Kalajdži Romani of Montana (Bulgaria). Besides the widespread *-as*, this dialect shows the previously unnoticed remoteness markers *-asa* and *-asta*, which have the same distribution of *-as*, but a different origin. The objective of the study is to propose a reconstruction of the genesis of the two variants. While the remoteness marker *-asa* can be explained as the outcome of recent internal innovation of Kalajdži, the remoteness marker *-asta* seems to be connected to the OIA root *sthā-* and, pointing to a more ancient phase of the language, suggests a higher complexity of the Proto-Romani strategies to build the imperfect and the pluperfect.

**Keywords:** Romani; Kalajdži; remoteness marker; imperfect; historical morphology; synchronic variation

### 1. Romani remoteness marker: a brief state of the art

Romani imperfect and pluperfect are built by the agglutination of the same morpheme to the inflected forms of the present and the perfect, respectively<sup>1</sup>. The marker used to build the imperfect and the pluperfect is labelled as “remoteness marker” (Matras 2001: 35) by the literature on Romani varieties. For example:

- (1) present *ker-av(a)* ‘I do’ → imperfect *ker-av-as* ‘do-1.SG-REM ‘I was doing/I did/I used to do’; perfect *ker-d-om* ‘do-PFV-1.SG’ ‘I did’ → *ker-d-om-as* ‘do-PFV-1.SG-REM ‘I had done’.

From a functional and a formal point of view, the remoteness marker is quite homogeneous in Romani varieties. It conveys a temporal value of distance towards a determined point of reference excluding at the same time any overlapping with the moment of speech<sup>2</sup> and thus its semantic spectrum seems very similar to one of the markers of “temporal discontinuity” highlighted by Plungian and Van der Auwera (2006). The main recorded forms in the different dialects are *-as/-a/-e/-s/-ys/-ahi* (cf. Matras 2002: 152) and in some dialects also *-sa(s)*, which in origin, and still in some varieties, might be an allomorph of *-as*. These forms allow to reconstruct a single Proto-Romani form *\*asi* (cf. Bloch 1932, Bubenik 1995) or, according to a recent proposal, *\*sasi* (Scala 2020). These reconstructed forms go back to the Old Indo-Aryan copula *as-* ‘to be’, maybe through the third person Middle Indo-Aryan form *āsī* or *āsī* ‘he/she/it was’ and, in the case of *\*sasi*, with an analogical alignment to the base *s-* which occurs in the other forms of the verb ‘to be’ (cf. Scala 2020: 238). Imperfect and pluperfect are thus the outcome of the grammaticalization of the third person past of the copula, which has been reanalysed as remoteness marker.

Even though the overall panorama is homogeneous, in some Romani dialects the remoteness morpheme has been renewed and its function is covered by a form which goes back to *sine*, a variant of the past third person copula, which can show, due to the grammaticalization process, a different degree of phonetic erosion, see for instance Arli *asala-hine* ‘he was smiling’ (<*\*asala sine*, cf. Boretzky 1996: 22), *dela hine* ‘he was giving’ (<*\*dela sine* cf. Boretzky, Cech, Iгла 2008: 29), Abruzzian Romani *kerásəna* ‘I was/we were doing’ (cf. Soravia 1977: 87, Scala 2020: 226)<sup>3</sup>. Thus, as underlined by Matras and Elšik (2006: 192), “few dialects present additional evidence for a greater complexity of the imperfect”. Besides, even though the uniformity of Romani remoteness marker in the documented varieties allows to project back a similar state of affair to Proto-Romani, nonetheless some dialects show noteworthy allomorphs and a certain level of internal variation that could suggest that this levelling has been preceded by a more complex situation. The following paragraphs will deal with the remoteness system of Kalajdži of Montana, characterized by different free variants which reflect both renewal in the system and what seems to be a trace of the aforementioned Proto-Romani complexity.

## 2. The remoteness markers of Kalajdži

The Kalajdži dialect documented by the Romani Morpho-Syntax Database (from now on, RMS Database)<sup>4</sup> under the label of BG-016 is a variety spoken in the Bulgarian region of Montana, near the town of Bergovica, in the Northwest of the country. The metadata in the Database report that the speakers are not migrant and that, although showing bilingualism with Bulgarian during the recording, they were in contact with Turkish in recent times, and with Greek in an earlier period.

Montana Kalajdži displays some features which are typical of the North Balkan dialects (following the label used by Elšik & Beníšek 2020: 401)<sup>5</sup>.

The section of the RMS Database which accounts for the verbal inflection of Montana Kalajdži (subsection Tense&Mood) reports the remoteness marker *-as* only, which, as previously mentioned, is the more common in the documented varieties. Nonetheless, an analysis of the sample in the whole questionnaire allows to identify three different remoteness morphemes, which are *-as*, *-asa* and *-asta*. Here follow three tables reporting a list of sentences for each marker; in each table, the first column contains the identification number assigned in the RMS Database to the sentence; in the second column I have extracted and glossed the verbs with the remoteness marker and in the third column I reproduced the whole sentence in which the verbs are found:

(2) remoteness morpheme *-as*:

405	<i>phir-áv-as</i> walk-1.SG-REM	<i>kana vujóm cikonoró često phiráv-as ko pazári</i> ‘When I was young, I used to go to the market very often’
411	<i>a-én-as</i> come.PFV-2/3.PL-REM	<i>áko aénas araci može bi štěše te dikhé<sup>6</sup> la</i> ‘If you had come yesterday, you would have seen her’
628	<i>mang-óm-as</i> want-1.SG.PFV-REM	<i>vinagi mangómas te žav indija</i> ‘I have always wanted to go to India’
674	<i>beš-ál-as</i> sit-3.SG-REM	<i>ek cikno žukél bešálas ži ko kher</i> ‘A little puppy was sitting beside the house’
960	<i>thov-él-as</i> wash-3.SG-REM ( <i>zéb-el-asa</i> sing-3.SG-REM)	<i>voj thovélas o parcája em zébelasa</i> ‘she was washing the laundry and she was singing’

In the last example, (2) 960, next to the form *thovélas* ‘she was washing’ showing the remoteness marker *-as*, we also find at short distance the form *zébelasa* ‘she was singing’, which displays the remoteness marker *-asa*. The use of this last marker is exemplified in the following table:

(3) remoteness morpheme *-asa*:

392	<i>arakh-él-asa</i> find-3.SG-REM ( <i>ov-él-as</i> be-3.SG-REM)	<i>áko arakhélasa o kher vov naj te ovélas akána katé</i> ‘If he had found the house, he wouldn’t be here now’
837	<i>zán-l-asa</i> know-3.SG-REM	<i>káva manúš žánlasa sar te opravizel e instruménte</i> ‘This man knew how to repair the instruments’
997	<i>ród-éj-asa</i> search-2.SG-REM	<i>arakhén i čánta káte ródejasa?</i> ‘Did you find the bag you were looking for?’
1018	<i>prənžan-áv-asa</i> know-1.SG-REM ( <i>khél-l-asta</i> dance-3.SG-REM)	<i>prənžanávasa eké čha káte khéllasta parángo</i> ‘I knew a young girl who used to dance for money’

In this second table, the point (3) 1018 shows the use of two different remoteness markers in the same sentence, i.e. the remoteness marker *-asa* and the remoteness marker *-asta*, which is documented also in the following table:

(4) remoteness morpheme *-asta*:

647	<i>phir-l-asta</i> walk-3.SG-REM	voj <i>phirlasta pála eké mrušúste</i> ‘She was walking behind a man’
686	<i>av-él-asta</i> come-3.SG-REM	<i>šunzóm múzika, voj avélasta tar o kher</i> ‘I heard music which came from the house’
905	<i>moiz-el-asta</i> can-3.SG-REM	<i>vov vuló kici silnó sóske moizelasta te banzerál o sástr</i> ‘He was so strong that he could bend the iron’
980	<i>prənžan-áv-asta</i> know-1.SG-REM	<i>prənžanávasta eké čha, la nisar na vulé la paráde láte</i> ‘I knew a girl, who never had money with her’
1027	<i>baró-l-asta</i> grow-3.SG-REM	<i>i péjka ándi kúxnnja kerdé la tar o kaš káte barólasta angál amáro kher</i> ‘The bench in the kitchen is made of the tree that was growing in front of our house’

From a functional point of view, the distribution of the markers *-as*, *-asa* and *-asta* does not display any observable motivation: the samples do not point at some semantic specificity of the affixes, nor a complementary distribution is observable in the paradigm. For example, the 1SG forms (4) 980 *prənžanávasta* ‘I knew’ and (3) 1018 *prənžanávasa* ‘I knew’ are used in a very similar context; the same stands for the optative/irrealis meaning, which characterizes the remoteness forms in (2) 411 *aénas* ‘[If] you had come’ and in (3) 392 *arakhélasa* ‘[If] he had found’. Two different markers are also used in the same function and in the same sentence, as in (2) 960 *thovélas* ‘she was washing’ and *zévelasa* ‘she was singing’ and in (3) 1018 *prənžanávasa* ‘I knew’ and *khéllasta* ‘she danced’<sup>7</sup>.

If the remoteness marker *-as* is a well-documented and frequent morpheme in the Romani panorama, the other markers, which are likely free variants of the marker *-as*, need some discussion.

A preliminary issue must be brought to light. The verbal system of this variety of Kalajdži opposes a present inflection to a subjunctive or subordinative inflection, which differ from each other just by a final *-a*, e.g. pres. *žáva* ‘I go’ vs. sub. *žav* ‘(that) I go’, pres. *žála* ‘he goes’, vs. sub. *žal* ‘(that) he goes’, pres. *mangáva* ‘I want’ vs. sub. *mangáv* ‘(that) I want’, pres. *mangéla* ‘he wants’ vs. sub. *mangél* ‘(that) he wants’. These forms with and without final *-a* will be referred to as “long” and “short” forms respectively (cf. Elšik 2020: 160). As we will discuss later, the value of present indicative was originally assigned to the short one; nonetheless, given the fact that the present of Kalajdži ends now in *-a*, one can wonder how to analyse the previously cited imperfect and pluperfect forms on a synchronic level, i.e. whether the first *-a-* of the remoteness morphemes *-as*, *-asa*, *-asta* is to be considered as part of the remoteness morphemes or as part of the long form to which they attach (e.g. *beš-ál-as* vs. *beš-ála-s*, *prənžan-áv-asa* vs. *prənžan-áva-sa*). The marker *-as* is recorded also in the pluperfect form (1) 628 *mang-óm-as*, built from the 1SG

*mangóm*, suggesting that the vowel *-a-* is part of remoteness morpheme *-as*, at least on an abstract level. For the other markers, the morphological rule is not clear, since we have no pluperfect forms recorded. Keeping in mind that the synchronic rule may be different, for the sake of mere graphic simplicity we will use the forms *-asa* and *-asta* and not *-(a)sa* and *-(a)sta*.

In the following paragraphs, we will deal with the variant *-asa* and with the variant *-asta*, and their possible sources will be discussed.

### 3. The remoteness marker *-asa*

The remoteness marker *-asa* found in *prənzánávasa* ‘I knew’, *zébélasa* ‘he was singing’, shows phonetical proximity both to the remoteness marker *-as* and to the remoteness marker *-sa(s)* (from an older *\*sas*) witnessed in the imperfect and pluperfect inflection of some dialects, as Abruzzian Romani, e.g. *kerén-sa* ‘you.PL/they were doing’, and as Ajia Varvara Romani, e.g. *peló-sas* ‘he had fallen’<sup>8</sup>. If *-asa* is to be linked to *-sa*, thus the segmentation of a form like *zébélasa*, at least from a diachronic point of view, should be *zébél-a-sa*, and thus we should account for the first *-a-* of *-asa*; if we traced back the morpheme *-asa* to *-as*, we should segment the form *zébél-as-a*, and thus the genesis of the final *-a* should be explained.

If we accepted that the Kalajdži forms in *-asa* had to be linked to the remoteness marker *-sa(s)*, the remoteness marker *-sa* (<*\*sasi*) should have been agglutinated to a form of the present ending in *-a*, namely to a long form of the present. Such an interpretation raises some questions about the internal chronology of the changes affecting the long or short forms of the present and the grammaticalization of the remoteness markers. The so-called long forms are in fact the outcome of a process of grammaticalization which has led the agglutination of a marker *\*-a* to the short forms, which, with an exception due to later restructuring, regularly proceed from the Old Indo-Aryan present inflection (Beníšek 2020: 33). These “new” long forms have the meaning of indicative present only in some Romani dialects<sup>9</sup>. The spread of the morpheme *\*-a* in the historical varieties testifies that it was in the Romani system during its common phase (usually called Proto-Romani or Late Proto-Romani) and that its grammaticalization process must have had its start in that period. Nonetheless, the different outcomes recorded in the dialects suggest that this change was surely not completed at the time of the diaspora which the Romani-speaking community went through in medieval times (cf. Beníšek 2020: 18) and which led to the current dialectal differentiation. Besides, due to the interdialectal variation of the function of the long forms, the etymology of *\*-a* is still unclear and the discussion about its original meaning is ongoing (cf. Beníšek 2020: 36)<sup>10</sup>.

As far as the imperfect and pluperfect forms concerns, since the remoteness marker *-as*, as we said, is functionally uniform and formally quite stable in Romani, its grammaticalization reasonably occurred and reached its stability in the common phase of Proto-Romani and therefore there is no doubt that it somewhat preceded the grammaticalization of *\*-a*. If *-asa* has to be interpreted as the agglutination of *-sa* to a long present, it follows that two waves of grammaticalization of the remoteness marker must have occurred: the first (*-as*) before the long present spread in the system,

the second (-*sa*) after this change. Nonetheless, a recent contribution by Scala (2020) convincingly dates the formation of the variant -*sa*, found in Abruzzian Romani and in Ajia Varvara Romani, to the same phase which led to the more frequent marker -*as*.

Scala's point of departure is the fact that the dialect of Ajia Varvara (described by Birgit Iгла in 1996) shows two allomorphs of the remoteness marker, namely -*as* and -*sas* (the second is also the third person past of the copula). The allomorph -*as* is post-consonantal, the allomorph -*sas* is post-vocalic. This distribution is well observable in the inflection of the pluperfect, which, as previously mentioned, is built by adding the remoteness marker to the forms of the perfect. The perfect has in fact vocalic ending for the third person of intransitive verbs (which are participles in origin), whether the first and second person of every verb and third person of transitive verbs have consonantal ending. So, in Ajia Varvara Romani we have a first person singular pluperfect *ker-d-ém-as* 'do-PFV-1.SG-REM' 'I had done', *pe-l-ém-as* 'fall-PFV-1.SG-REM' 'I had fallen' but a third person singular pluperfect *ker-d-ás-as* 'do-PFV-3.SG-REM' 'he had done', *pe-l-ó-sas* 'do-PFV-M.SG-REM' 'he had fallen'. According to Scala (2020: 233-234), the third person perfect of intransitive verbs (namely a form like *pe-l-ó-sas*) could have been the starting point of the reanalysis which led the copula to be interpreted as a marker of tense. The third person perfect of intransitive verbs coincides in fact with the past participle, so the form *peló* may mean both 'he felt' and 'fallen (masculine singular)'. The double function of this form could have generated the shift from *peló sas* 'he was fallen' to *pelósas* 'he had fallen'. If we assume a form like *sas* as the starting point for the grammaticalization that led to the remoteness marker, the allomorphy -*as/-sas*, found in Ajia Varvara Romani, may be caused by a reanalysis of the third person singular forms of the pluperfect of the transitive verbs, e.g. \**kerdás-sas*, which is functionally and structurally equivalent to forms like *peló-sas*. In fact, Romani does not have phonological consonantal length, and a form like *kerd-ás-sas*, built in analogy with *peló-sas*, could have been reanalysed as *kerd-ás-as* allowing the creation of the allomorph -*as* as a post-consonantal variant. From such forms, the remoteness morpheme could have been extracted and extended at first to the other persons of the perfect inflection to form the pluperfect (starting from the third person perfect of transitive verbs, which in some varieties shows an oscillation between a participial form like *kerdó* and the form *kerdás*), then to the present, forming thus the imperfect. The diffusion of the marker -*as* in most of the Romani varieties must be caused by the fact that the ancient inflection of the present is the one with short forms, thus with final consonants. Another point in favour of an origin of the remoteness marker from the reanalysis of a copula like *sas*, i.e. with initial *s-*, is the fact that the third person past copula *sas*, preserved mainly in Vlax varieties (cf. Gjerdman & Ljungberg 1963: 121; Boretzky 1994: 60), is quite ancient, since it is found also in other branches of Romani which are spoken in different and distant areas of the Romani panorama, for instance in East Slovak Romani (RMS Database SK-002), East Finnish Romani (RMS Database FIN-002) and in Welsh Romani (Sampson 1926: 209). Based on these observations, Scala (2020: 238) proposes thus a new etymology for the allomorphic variants -*as/-sas*, which is the Proto-Romani

*\*sasi* which may account for the variation of the remoteness morphemes. This form originated from the Middle Indo-Aryan form *\*asi* ‘was’, which, in analogy with the other forms of the copula, acquired an initial *s-* (well witnessed in the copula of the historical varieties). If this is true, the hypothesis that Montana Kalajdži *-asa* was built from a long form of the present plus the remoteness marker *-sa* is not plausible. It is not consistent with the data we have from the other varieties, which show that *-sa* likely grammaticalized before the spread of the long forms in *\*-a*. Besides, if we postulated two waves of grammaticalization to account for *-sa*, we should explain why such second change would have started, since we can easily document the stability of the remoteness marker *-as*. The reconstructions proposed up to now, instead, more economically place the start of the allomorphic pairs of the remoteness marker to the same phase.

The second and most likely way to explain the forms in *-asa* is that they were *-as* forms, thus built by adding the marker *-as* to the short present, but they went through the agglutination of another formant as a result of an internal innovation of this dialect. This leads us to examine what the final *-a* of the marker could be and why it was added.

We must say in fact that some Romani varieties, such as for instance Vlax, Southern Central and South Balkan varieties (cf. Boretzky & Iglá 2004, Teil 2: Karten 17-21), show a change of etymological final *-s*, which goes through aspiration (> *-h*) or deletion. In Northern, Central and in some Balkan Romani dialects, the initial and intervocalic *s* as well may go through aspiration (> *-h*), jotation or deletion, even though in these cases the alternation *s/h* may be in part inherited and it occurs only in morphologically determined contexts, as for instance the second person singular of the long forms, e.g. Kalajdži *mangéja* ‘you want’ < *\*mangesa*. A detailed study of the *s/h* alternation in Romani was conducted by Matras 1999.

As far as final *-s* concerns, in Kalajdži we have regular deletion in all the contexts listed by Matras except in the remoteness marker *-as*<sup>11</sup>. It is thus likely that the remoteness marker was partially preserved by the general deletion of *-s* because of a specific constraint: if the change was extended to the remoteness marker, the imperfect would have merged with the present, which, as we saw, has the long form (i.e. *keravas* ‘I was doing’ > *\*kerava* ‘I do/I was doing’). It is also likely that, besides the constraint, the imperfect forms were redetermined with further morphological material in order to avoid the final position of *-s*, which was phonetically weak. A possible source for the morphological material is the remoteness marker *-as* itself, which may have been reduplicated in order to avoid the apocope of final *-s*, later occurred to the added morpheme (e.g. *keravas* > *\*kerava(s)* > *\*keravasas* > *keravasa*).

The hypothesis of a reduplication of the remoteness morpheme *-as* is not the only one which may be worth of attention. A second explanation seems even more likely: the possible source of the final *-a* of the marker *-asa* may be the result of paradigmatic pressure exerted by the long forms.

In Kalajdži, as previously mentioned, the verbal system opposes a long present indicative (*keráva* ‘I do’, *mangáva* ‘I want’) to a short subjunctive (*keráv*

‘[that] I do’, *mangáv* ‘[that] I want’). There is no sign of oscillation between long/short forms and so this opposition has acquired a modal value. The imperfect in *-asa* could thus be the result of the analogical extension of the *-a* of the present to the forms of the imperfect, which also share with the present some aspectual features, e.g. they are both imperfective.

An extension of *-a* triggered by paradigmatic pressure seems to occur also in other varieties, for instance in Mečkari (RMS Database AL-001), in which we have present long forms, subjunctive short forms and in which the *-a* is sometimes extended to the imperfect, e.g. *khelélas-a* ‘she was dancing’ (1018), *našlomas-a* ‘I had gone’ (398), *phirélas-a* ‘he was walking’ (461), and, interestingly, to the past inflection of the copula, e.g. *(i)sinom-a* ‘I was’, *(i)sinan-a* ‘you were’, but it does not join the forms of the perfect.

#### 4. The remoteness marker *-asta*

The other variant of the remoteness marker attested in Kalajdži is *-asta*. This morpheme is remarkable since, to my knowledge, it was never reported in grammars or scientific literature up to now.

A similar marker, nonetheless, is documented in some written sources which hitherto escaped the attention of the scholars. The texts at issue are written in two Sinti dialects, i.e. varieties of the North-western group of Romani dialects, spoken in northern and central Italy<sup>12</sup>.

Regarding the varieties spoken in northern Italy, we have some documents in a dialect that can be considered as Lombard Sinti. Such sources are three short tales composed by the same author and published in the journal “Rom. In cammino” from 1978 to 1981, and some poems and short tales composed by young speakers in Pontelagoscuro (in the province of Ferrara, on the borders of Veneto Region) and published by Santino Spinelli in 1995 and 1996<sup>13</sup>.

The following text shows some examples of imperfect forms from these texts in Lombard Sinti<sup>14</sup>:

- (5) *tinkareato ar č-aj-asta misto kuantu n’ i-s-mi kaja perla*  
 think how stay-1PL-REM well when not be-REM-1PL this pearl  
*pustarde; kam-aj-as-me, mengar mal kam-en-as-me [...].*  
 cursed; love-1PL-REM-Pron.1PL, our friend love-1PL-REM-Pron.1PL  
*Mengur čao s-al-asta, gjav-el-asta, i-s-lo perdo da ger,*  
 our son laugh-3SG-REM, sing-3SG-REM, be.3-REM-M.SG full of joy  
*akana na prindžarajalo butar.*  
 now not recognize.him anymore.  
 ‘Do you remember how well we were when we didn’t have that cursed earl:  
 we loved each other, our friends loved us [...]. Our son laughed, sang, he  
 was full of joy, now we don’t recognize him anymore’ (Oliviero 1981)



In these documents, we find both *-as* and *-asta* as remoteness morphemes. As in the case of Kalajdži, it is not clear whether the use of one marker or the other has some motivation and *-as* and *-asta* rather appear as free variants.

The other Sinti variety which records a morpheme that can be linked to *-asta* is Shinto Rosengro, an old settlement Sinti dialect discovered by Leonardo Piasere in the late 90s' (Piasere 1996, Caccini, Barontini, Piasere 2001) and once spoken in central Italy. Shinto Rosengro is documented by some manuscripts dated from the 1892 to 1912, now stored in the Biblioteca Planettiana of Jesi, in Central Italy, and in the Biblioteca Teresiana in Mantova. The author of the manuscripts was Sigismondo Caccini, a man who lived with the Shinte Rosengre and took part to their short-range travels through Central Italy. The documents he wrote, which have been analysed in my Phd Thesis (Meli 2019), show a Sinti variety carrying some peculiar innovations, which distinguish it from the other old settlement varieties recorded up to now on the Italian ground (Sinti and Southern Italy Romani dialects such as Abruzzian Romani). Unluckily, we don't have traces of currently living speakers and thus Shinto Rosengro seems extinct now.

The only remoteness marker recorded in Shinto Rosengro is *-esta*, likely from an older *-asta*. The outcome of *-e-* from a former *-a-* is not new among Sinti varieties: the anteriorization *-as* > *-es* in the context of the remoteness marker can be found, for example, in the Piedmontese Sinti short tales collection *O ker kun le penjá* 'The house with the wheels' by Annibale Niemen (1995), which usually shows forms like the 3SG imperfect *dél-es* 'he did' (*d-él-es*, 'give-3SG-REM') instead of the more widespread *dél-as*; a remoteness marker *-as*, with a central vowel *ə* instead of *a*, is also witnessed by other Sinti varieties, and it is explained by Norbert Boretzky (1995: 23) as the outcome of the reduction of the unstressed vowel due to the contact with German<sup>15</sup>. The *-a-*, as we will see, is still preserved in Shinto Rosengro in the third person singular of the copula.

The rich amount of data from this variety helps us to document for this marker the same distribution and function of the remoteness markers attested in the other Romani dialects. In the following table some examples of Shinto Rosengro remoteness marker *-esta* are presented<sup>16</sup>:

(6)

a.	<i>gi-ass-esta</i> 'go-1PL-REM' 'we usually went'  <i>d-en-esta</i> 'give-2/3PL-REM' 'they usually gave'	<i>Por Pierlati na giassesta buttidir for devleski, oski e raye Ferretti [...]</i> <i>denesta devléski ki mólo kai bersh</i> 'In Pierli, we <b>didn't go</b> begging anymore because Mr. Ferretti [...] <b>gave</b> charity once in a year'
b.	<i>pir-esta</i> 'walk-REM' 'she was walking'  <i>v-esta</i> 'come-REM' 'she was coming'	<i>Gorda pirsta kajardapi ka pelal late vesta yek romni klisti</i> 'while she <b>was walking</b> , she noticed that beside her a woman <b>was coming</b> , riding a horse'

c.	<i>pi-áv-este-lo</i> 'drink-1SG-REM-Pron' 'I would drink it'	<i>piávestelo ger</i> 'I would drink it with pleasure'/'I would gladly drink it'.
d.	<i>nak-i-esta</i> 'pass-PF-REM' 'she had passed'  <i>mangh-i-esta</i> 'ask-PF-REM' 'she had asked'  <i>d-esta</i> 'give-REM' 'he used to give'	<i>Bute mole nakiesta tilal lengro ker ki puri shinti ke manghiesta devleski kai puri bibi, mek doadoi puri neidighi na desta sarfar ci</i> 'Many times an old shinti <b>had passed</b> under their home <b>asking</b> charity to the old aunt, but that old miser never <b>gave</b> anything'
e.	<i>l-i-en-esta</i> 'take-PF-3PL-REM' 'they had taken'	<i>dui romes lienesta for romiake duien pegná</i> 'two rom/men <b>had taken</b> two sisters as their wives'
f.	<i>s-asta</i> 'be-REM' 'it was'  <i>gia-sta</i> 'go-REM' 'she was going'	<i>Sasta i rat kristuneskero ki tuvani giasta pirangani kangriate</i> 'It was Christmas night, and a lady farmer <b>was going</b> to the church by feet'
g.	<i>s-en-esta</i> 'be-3PL-REM' 'they were'	<i>Senesta etske</i> 'they <b>were</b> witches'

The third person singular of the imperfect is built by adding the remoteness marker to the imperfective stem of the verb, thus with a zero realization of the person marker. This, in all likelihood, is a result of a phonetic reduction, e.g. *kerélesta* > *kerélsta* > *kerésta*. The zero realization of the third person singular extends by analogy to the third person singular of the pluperfect, which is built by adding the marker *-esta* to the perfective stem of the verb, showing thus the paradigmatic solidarity of the third person singular through the verbal paradigm<sup>17</sup>. The absence of a phonetical realization of the third person is consistent with Benveniste's remarks on the different status of the third person in comparison with the one of the first and second person (Benveniste 1946: 4-6). In his view, the third person is interpreted as a 'non-person' because it lacks individuability and specificity, which are inherent properties of the first and the second person: the third person is mainly anaphorical and may or may not be referred to an element in the situation shared by the speakers, while the first and second person are inherently deictic and must carry the reference to such situation. Such difference in the core semantic properties of the third person in comparison with the first and the second may have also morphological expression and result in the frequent zero marking of the third person, which is observable also in Shinto Rosengro.

The functions of the verbal forms exemplified in the table coincide with the one attested in the other Romani varieties, encompassing the optative/desiderative and conditional/irrealis meaning, as shown in (6)c *pjavestelo* ‘I would drink it’.

The presence of the morpheme *-asta/-esta* in varieties so far from each other, seems to point at an ancient origin and it’s more reasonably due to some inherited form than to the internal evolution of each variety.

Due to its rareness, the remoteness morpheme *-asta* was never noticed nor accounted for and its etymology did not receive any attention so far. In the following paragraphs, I will propose an etymology through the analysis of two possible ways to explain the origin of the marker.

#### 4.1. Seeking for the etymology of *-asta*: grammaticalization of the morpheme *-tar*?

If we consider the similarity between the remoteness marker *-as* and the first part of *-asta* and cross this fact with the homogeneity of the marker *-as* in the Romani varieties, we may suppose that the morpheme *-asta* may be the outcome of a grammaticalization process which involved the agglutination of some marker to *-as*.

Due to its phonological proximity and its occurrence with verbal forms, the particle that could be taken into consideration as source for *-ta* is the morpheme *-tar*; that we find for instance in Vlach *gelo-tar* ‘he went away’.

The morpheme *-tar* is an unstressed affix used with verbs. The marker is formally identical to the morpheme of the pronominal and nominal ablative case, that usually expresses the origin or source of a movement (e.g. *kerés-tar* ‘from home’, *lés-tar* ‘from him’). The verbal suffix *-tar* is common in Vlach varieties, but it is attested also in non-Vlach dialects in which an inter-dialectal borrowing can be excluded (cf. Boretzky & Iglá 2004, Teil 1: 174). Its functions are not homogeneous in Romani, as underlined by Norbert Boretzky and Birgit Iglá in their *Atlas*. According to Ian Hancock (1995: 100), in Vlach varieties *-tar* is added to the inflected verbs of motion to express the meaning ‘off’, ‘away’, and it is considered by the author as a different morpheme from the *-tar* of the nominal inflection. Hancock illustrates the use of verbal *-tar* with examples like *te teliarástar akana!* ‘let’s set off now!’ or *našélastar lestar* ‘she was running away from him’ (Hancock 1995: 101). Almost a century before Hancock, Alexandre Paspatis, in his account of the Romani dialects of Thrace, proposed another interpretation. In the Balkan varieties that he describes, which are not Vlach, the verbal *-tar* is used only with the participial forms, especially with the participial third person of the perfect<sup>18</sup>, and it indicates that the predication of the verb has had an end, namely it has a resultative or perfective meaning: “on ajoute à la fin des participes, la syllabe *-tar*. [...] L’r final du *-tar* est fortement prononcé, de manière qu’on ne peut pas se tromper, et prendre *-tar* pour la conj. *ta*, ‘et’. *Tar*, en s’unissant au participe, représente l’action comme finie. Pour la plupart, il s’unit au participe, de la 3<sup>me</sup> pers. de l’aorist, soit au singulier soit au pluriel, *Aló ta beshtótar*, ‘il vint et il s’assit’. *Ghelótar yek tanéste*, ‘il alla dans un endroit’. *Pelótar ko khurdó pral*, ‘il tomba sur le frère cadet’” (Paspatis 1870: 100).

Even though further analysis should be needed, as an initial survey on the status of *-tar*, I checked some varieties of RMS Database, in order to enlighten the possible links with the remoteness marker *-asta* found in Kalajdži and Sinti. About 40 varieties were examined and two tendencies can be observed. Regarding the first tendency, we have varieties in which *-tar* is used just with participial forms of the perfect and with imperative. In this case, the morpheme *-tar* is not strongly related to the meaning of ‘away’ nor it occurs only with the verbs of movement. An example could be Spoitori (RMS Database RO-006): 801 *ou mulótar de serécie* ‘he died of poverty’, 971 *ou mulótar eke bolátar ne penžardi* ‘he **died** of a mysterious disease’, 589 *but mrušá avinétar kaj o bjáu* ‘many men **came** to the wedding’. These varieties show also *-tar* in the third person present of the copula, mainly with forms which have participial origin, namely the stem *sin-/in-*, as it is observable in the aforementioned Spoitori (RMS Database RO-006): 680 *angál i khangeri sinítar i škóla* ‘opposite the church **there is** the school’, 813 *kadá sinótar o čháó ki djóm les me báne* ‘this **is** the boy who I gave the money to’.

There are also some varieties in which *-tar* is used only with the verb *dža-* ‘to go’, and in these cases it means ‘away’, as in Kaldaraš (RO-008) 363 *phejále, žántar kathál!* ‘sisters, **go away** from here!’, or in Gurbet (HR-001) *phejálen džántar!* ‘sisters, **go away!**’. The morpheme *-tar* may be found also with *av-* ‘to come’ and other verbs of movement. In varieties which show this tendency, it is possible to find *-tar* with the imperfect, but this use is rare, and it seems the result of analogical extension.

Sinti varieties, as expected, do not show this morpheme, while in Kalajdži of Montana we find the verbal *-tar* just in the third person participial form, that is with the perfect and with the third person of the present copula *sino/ino*. For instance: 871 *voj inítar nasvai em xasála* ‘she **is** sick and coughs’, 416 *pósle panč minúte vov započnisílotar te kerél pherés* ‘after five minutes he **started** to talk’, 394 *voj trašávžitar kaná dikhé le* ‘she **became scared** when she saw him’. In this variety *-tar* combines very seldom to verbs of motion, mainly with the verb *av-* ‘come’, and only with the participial forms, e.g. 748 *nékakvo si bogáto mruš alótar* ‘some rich man **arrived**’, 465 *voj pánda na aítar ki škólja* ‘she has not **come back** to school yet’. The distribution of *-tar* in Kalajdži of Montana seems thus more similar to the one found in the varieties described by Paspati 1870.

This initial survey seems to suggest that *-tar* might have extended from participial forms to the other verbal forms, likely reaching at first the imperative than the other inflected forms (subjunctive and present tense) by analogy.

Drawing some conclusions about the possible origin of the ending *-ta* of the remoteness marker *-asta*, we can affirm that, irrespective of their phonetical similarity, the distribution of *-tar* is not very compatible with the picture we should imagine as a source for the second part of the marker *-asta*.

As a matter of fact, if we postulate such origin for *-ta*, we must imagine that *-tar* went through grammaticalization and merged to a past form of the verb ‘to be’. Nonetheless, in the historical dialects this marker is never used in combination with the remoteness marker *-as-*. Besides, from a semantic point of view, both the meanings

of *-tar*, i.e. the value of ‘away from’ and the resultative or punctual meaning, seem incompatible with imperfect and remoteness marker, which is, instead, a marker of durativity.

#### 4.2. Another etymological path: a new copula for the Proto-Romani

Once the hypothesis that *-asta* is *-as* + *-tar* has been rejected, a new etymological path must be followed to explain this morpheme.

The presence of an identical marker in varieties, such as Kalajdži and Sinti, which are quite distant from each other both from a geographical and a structural point of view, leads us to look for an origin that is chronologically located before the European diaspora.

To sum up, in Romani we have two tenses, imperfect and pluperfect, which are built by adding, to the inflected form of the present and perfect respectively, a marker that goes back to a copula form, probably with the meaning ‘was’. The different remoteness markers in Romani were reconnected to the same copula *\*(s)asi* (< MIA *āsi* or *āsī* ‘he/she/it was’), which doesn’t match with the marker *-asta/-esta* and which cannot be its ancestor.

Since the marker seems to be ancient, we may be in front of a relic of another variant of a grammaticalized copula. Some hint could come from Domari, a New Indo-Aryan variety which in many ways is the nearest to Romani among the New Indo-Aryan dialects. In the literature of XIX and early XX century about the so called “Syrian Gypsies” (the speakers of Domari), the third person copula is *asti* in the present, and *asta*, *āšti* or *āšta* in the past.

For instance, John Sampson (1926) displays the paradigms of Welsh and Greek Romani copula in comparison with the paradigms of Syrian Gypsy copula found in Pott’s *Ueber die Sprache der Zigeuner in Syrien* (1845) and in Macalister’s *The Language of the Nawar of Zutt, the Nomad Smiths of Palestine* (1914). The following table reproduces the forms given by Sampson (1926: 208, 209; in bold the third person copula mentioned before):

(7)

		Present			Past			
		Welsh Romani	Greek Romani	Domari	Welsh Romani	Greek Romani	Domari	
				Pott			Pott	Macalister
SG	1	<i>īšóm, śom</i>	<i>isóm</i>	<i>stūmi</i>	<i>somas</i>	<i>isómas</i>	<i>stūma</i>	<i>āštōm(i)</i>
	2	<i>san</i>	<i>isán</i>	<i>stūri</i>	<i>sanas</i>	<i>isánas</i>	<i>stūra</i>	<i>āštūr(i)</i>
	3	<i>sī</i>	<i>isí</i>	<b><i>asti</i></b>	<i>sas</i>	<i>isás</i>	<b><i>asta</i></b>	<b><i>āšti, āšta</i></b>
PL	1	<i>īśám, sam</i>	<i>isám</i>	<i>steini</i>	<i>samas</i>	<i>isámas</i>	<i>steina</i>	<i>āštēn(i)</i>
	2	<i>sen</i>	<i>isán</i>	<i>steisi</i>	<i>senas</i>	<i>isánas</i>	<i>steisa</i>	<i>āštēs</i>
	3	<i>sī</i>	<i>isí</i>	<i>steindi</i>	<i>sas</i>	<i>isás</i>	<i>steinda</i>	<i>āšte</i>

Noteworthy are also the data shared by Kerope Patkanoff (1908). In a series of articles on the Journal of the Gypsy Lore Society he collected and documented some data of the Karaći, a variety of the so-called northern Domari, which shows, for the present copula, the following forms:

- (8) *astum* ‘I am’,  
*astoj* ‘you are’  
*astaq, asta, a, aj, haj, hi*, ‘he is’ (Patkanoff 1908: 265)

The form *asta* is thus attested in Karaći as a present copula. The root *st-/št-*, observable in the data from (7) and (8), is still used for the copula of the northern Domari varieties, which are uniform in this respect.

If we consider more recent sources, and especially the description of the Domari of Aleppo (a northern variety) by Bruno Herin (2012), and the work of Yaron Matras (2012) on the Domari of Jerusalem (a southern variety), we also find a fixed grammaticalized form *ašti* (present) ‘there is’ and *ašta(ši)* (past) ‘there was’ used in existential constructions and in possessive construction. This last form is related from scholars to the OIA *sthā-* ‘to stand’. (Matras 2012: 265).

Therefore, as far as the etymology of the Romani remoteness marker *-asta/-esta* concerns, an account for our data could be that, next to *\*(s)asi*, another copula form, *\*asta*, could have developed, allegedly from the OIA *sthā-* ‘stand’, for instance from the OIA aorist *āsthāt* (cf. Whitney 1896 § 830). Such a form could match from a semantic perspective, since it is a third person copula of the aorist, and it could be a plausible precedent also from a phonetic point of view: both OIA *ā* and the OIA internal cluster *-st-* continue in Romani (cf. OIA *grāma* > *gav* ‘village’, OIA *svastha* > *sasto* ‘sane’; OIA *hasta* > *(v)ast* ‘hand’). The copula form *\*asta* could have followed the same grammaticalization process as *\*(s)asi* but remaining just in a small number of varieties. Relic of this form could be the remoteness marker *-asta* found in Kalajdži and in some Sinti varieties, and in the imperfect third person of Shinto Rosengro copula *sasta* ‘he was’. No further traces of that were founded in the varieties I checked up to now.

The parallel between Domari and Romani could be not so inconclusive, because even though there is no strict proof that these languages ever were a single unity, nonetheless they do share a series of innovations isolating themselves from the other NIA, and also convergent developments showing that they shared the same geolinguistic ground in different periods (Matras 2012: 20-27).

Furthermore, the possibility for a same variety to have different auxiliaries, with or without a complementary distribution, is not unknown to Domari and Romani: for the existential predication we have in Romani the verbs *som* (mutatis mutandis in the different varieties) going back to the OIA *as-* and the verb *ov-*, going back to the OIA *bhū-*, and shared with Domari, in which it goes along with *st-*, from the OIA *sthā-* ‘stand’.

Besides, the existence of different variants from the same person of the same auxiliary, i.e. a high level of polymorphism, is a quite frequent fact in the varieties

I examined (especially in the third person): even in the Kalajdži of Montana we have for the third person present *i, isi, inotar, (notar)*, while the past is built with *vul-* ‘become’. The same happens in Domari, as we saw in Karači (Patkanoff 1908), in which the third person singular of the present copula is expressed with many variants: *astag, asta, a, aj, hai*, and *hi* (which, just as a hint of secondary importance, are strikingly similar to Romani).

It could be not so surprising if, at the beginning of the grammaticalization process that led the third person imperfect of the copula to become a remoteness marker, a step has occurred in which the *\*sasi* or *\*asta* forms were both understood as auxiliary and fully interchangeable in the same context.

If we widen our gaze to other New Indo-Aryan varieties, we also find some interesting parallels regarding the use of the OIA root *sthā-* for new past constructions, possibly conveying a durative meaning. See, for instance, the following data (from Masica 1991, Grierson 1916 and Bubeník 1995):

(9)

Hindi		present	past	
	habitual	<i>ātā hū</i> (1SG.M) ‘I come’	<i>ātā thā</i> (1-3SG.M) ‘I/you/he used to come’	(imperfective participle + auxiliary)
	continuous	<i>ā rahā hū</i> (1SG) ‘I am coming’	<i>ā rahā thā</i> (1-3SG.M) ‘I was/you were/he was coming’	(short absolutive + perfective participle of <i>rahṇā</i> ‘stay’ + auxiliary)
	perfective	<i>āyā hū</i> (1SG) ‘I have come’	<i>āyā thā</i> (1-3SG.M) ‘I/you/he had come’	(perfective participle + auxiliary)  (cf. Masica 1991: 292)

Haryāṅvī <sup>19</sup>	imperfect	<i>māradā-thā</i> (1-3SG) ‘I/You/He was striking’ <i>marē-thā</i> (1-3SG) ‘I/You/He was striking’	(imperfective participle + auxiliary) (verbal noun + oblique marker <i>ē</i> + auxiliary)
Haryāṅvī of Rothak	imperfect	<i>mārai-thā</i> (3SG) ‘he was striking’	(present + auxiliary)  (cf. Grierson 1916: 255; Bubeník 1995: 9)

The Hindi *thā* (< OIA *sthā-*) is a verbal noun inflected by gender and number (femm.sing./pl. *thī*, masc.pl. *thē*). In Hindi it is an auxiliary used to mark the past tense. In fact, in past habitual, past continuous and past perfect, the tense is expressed by the opposition between two auxiliaries: the present is marked by the inflected present of the verb *honā* ‘to be, to become’, which is linked the OIA *as-* probably blended with OIA *bhū-* (cf. Oberlies 2005: 37), while the past is marked by *thā*. Similarly, in Haryāṅvī *thā* is postponed to the imperfective participle or to a verbal noun in the oblique case to build the imperfect. As already pointed out by Bubeník 1995, the form *thā* is also used in the Haryāṅvī of Rothak to build the imperfect. The imperfect of Rothak Haryāṅvī shows a construction quite similar to the Romani

imperfect, i.e. an inflected present plus a marker which conveys the meaning of imperfect. The form *thā* attested in these NIA varieties is inflected by gender and number and certainly proceeds from the OIA verbal adjective *\*sthitakah* (> *\*thiau* > hindi *thā*). However, it seems impossible to place such a verbal adjective as ancestor of Romani *-asta/-esta*, since OIA *-akah*, expressing gender and number in the OIA verbal adjective *\*sthitakah*, becomes always *-o* in Romani (cf. Beníšek 2020: 29) and the OIA initial *sth-* usually corresponds to Romani *th-* (cf. OIA *sthūlakah* > Romani *thulo* ‘fat’). For this reason, the hypothesis of an OIA aorist *āsthāt* as source of the Romani marker *-asta* remains more plausible<sup>20</sup>. If this hypothesis is correct, Romani would be in line with other NIA varieties which developed past tense markers based on the same OIA root *sthā-*.

## 5. Conclusions

The remoteness marker and the strategies to build the imperfect and the pluperfect are very uniform in Romani. Hence, the studies are unanimous in reconstructing a Proto-Romani past copula *\*(s)asi*, traced back to the OIA *as-* ‘to be’. This copula went through grammaticalization and acquired the function of remoteness marker, which was, and still is, agglutinated to the present to build the imperfect and to the perfect to build the pluperfect.

Kalajdži shows a certain degree of internal variation: we record three different remoteness marker which behave like free variants. Such variation witnesses both recent restructuring of the system and some hint that could enlarge our reconstruction of Proto-Romani, adding some details to the genesis of the remoteness maker in Romani.

In particular, we analysed two remoteness markers which were never mentioned by the literature to my knowledge, i.e. the marker *-asa* and the marker *-asta*. The first is likely the result of a recent internal innovation, triggered by analogy with the long form of the present, and/or by the phonetical changes which led this variety to lose the final *-s*.

The morpheme *-(a)sta*, instead, is witnessed also in varieties which are geographically and historically far from Kalajdži, is comparable with the copula of some Domari varieties, and thus has to be quite ancient, being in all likelihood etymologically linked to the OIA *sthā-*. The data led thus to the proposal of the coexistence in Proto-Romani of the reconstructed past copula *\*(s)asi* with another form for the copula, *\*asta*, which was likely involved in the same grammaticalization process that led to the more frequent marker *-as*. This hypothesis seems to find an interesting confirmation in some NIA varieties, which show markers of durative past surely deriving from a copula based on the OIA root *sthā-*.

## NOTES

<sup>1</sup> The label “perfect” indicates here a synthetic form built by adding the perfective personal concord markers to the perfective stem of the verb; in the literature on Romani varieties, the same tense/aspect is referred to also as aorist or preterite.



<sup>2</sup> In Matras' words, through the remoteness marker, "an event is contextualised relative to 'O' (=the 'origo'; cf. Reichenbach 1947, Bühler 1934). More precisely, remoteness places the event outside the reach of 'O' by excluding overlap between 'R' (= the contextual point of reference) and 'O'" (Matras 2002: 152).

<sup>3</sup> For a wide overview of the remoteness marker of Romani, cf. also Boretzky & Iglá 2004, Teil 2: Karte 137.

<sup>4</sup> The Romani Morpho-Syntax Database is a very rich online database which collects recordings, transcriptions, and descriptive sheets of 186 varieties of Romani spoken in the world (mainly in Europe). Since different examples from the RMS Database will be dealt with in the course of this study, to simplify, I will put the indication 'RMS Database' before each alphanumeric label of the varieties cited from the RMS Database, e.g. RMS Database BG-011.

<sup>5</sup> These dialects were previously referred to as South Balkan II (Boretzky 2000), Drindari–Kalajdži–Bugurdži (Matras 2002) and Balkan *zis*-dialects (from the name the speakers use to call the day, which is *zis* instead of *dives*, Elšík and Matras 2006); regarding their characteristic traits, these dialects typically show some innovative palatalisations, for instance the palatalization of the dental stops and the dental lateral [l] followed by a [i] or [j], e.g. *cikno* 'small' < *tikno* 'small', *šundzom* 'I heard' < *šundjom*, *zis* 'day' < *dives* 'day', *vuj* 'she was' < *vuli*, *ai* 'she came' > *ali*; they lose the *-d-* perfect in different verbs, e.g. *kerom* < *kerdjom*, they show the elision of the thematic vowel *-e-* in the present and imperfect of the verbs with stem ending in *-r-*, *-l-*, *-n-*, e.g. *kerla* 'he does' < *kerela*, *khella* 'he dances' < *khelela*, *šunla* 'he hears' < *šunela*, for an overview of the North Balkan traits, cf. Boretzky & Iglá 2004, Teil 1: 242-243 and Boretzky 2000.

<sup>6</sup> The form *štěše te dikhé* 'you would have seen' seems to be a partial calque of the Bulgarian future in the past (cf. Bulg., щеше да видиш 'you would have seen').

<sup>7</sup> The understanding of the distribution of these markers would certainly benefit from further studies based on a larger corpus than the one offered by the RMS Database: the RMS Database has an immense value and provides a rich amount of data and analysis otherwise inaccessible, but it is by necessity limited in scope and it is not designed to account for some dimensions of the language, such as for example the sociolinguistic domain.

<sup>8</sup> The dialect of Aija Varvara is part of the *vlox* group, which originated in the Romanian-speaking area and which is now the most widespread Romani dialect group in the world (cf. Elšík & Beníšek 2020: 405 and 406 for a brief overview of some diagnostic features of the group), while Abruzzian Romani is an isolated variety spoken in southern Italy which, according to Matras (2002: 10), appears as an "early offshoot of the Balkan dialects".

<sup>9</sup> The functions of the opposition long vs. short forms show interdialectal variation: in some dialects, as we saw for Kalajdži in §2, the long form is used for the present and the short for the subjunctive, but in other dialects the system opposes a short present to a long future (e.g. Lithuanian Romani, Tenser 2005: 29), or a default short present to a confirmative long present (e.g. Erli Romani, cf. Boretzky 1998: 141).

<sup>10</sup> While Bubeník (1995: 3-6) assumes a future source, according to Elšík & Matras (2006: 82-83) the morpheme *\*-a* had an indicative function opposed to the zero marking of the subjunctive; the more recent study on the topic (Scala, forthcoming), proposes an original progressive meaning.

<sup>11</sup> The contexts in which the final *-s* goes through deletion in Kalajdži are the masculine nominatives originally ending in *-os*, cf. Kalajdži *foro*, the masculine accusatives ending

in *-es*, e.g. Kalajdži acc. *dade* (< \**dades*), and the third person singular of the perfect of transitive verbs *-as*, e.g. *kere* (< \**kerjas*).

<sup>12</sup> The North-western dialects, such as Sinti-Manuš dialects and Finnish Romani, are spoken in western and north-eastern Europe; they share some innovative morphological features, such the extension of the third person singular *-as* of the transitive perfect to the intransitive one, which occur next to some archaic features, such as the preservation of the old second person singular of the present *-al*, substituted elsewhere by the ending *-an* in analogy with the second person plural; these dialect share also lexical traits, such as productive onomasiological genitives (e.g. *grajeng(er)o* ‘horse dealer’ from *graj* ‘horse’) and a German layer (cfr. Elšik & Beníšek 2020: 407-408, Boretzky & Iglá 2004: 277–289).

<sup>13</sup> These last documents were written in the occasion of a competition, named “Amico Rom”, which is still organised, and it has reached its XXVIII edition.

<sup>14</sup> Since it is not always recorded in the sources, this text and the ones in Shinto Rosengro do not display systematically the accent of the word.

<sup>15</sup> The evidences of an intense contact with German are numerous and common to all Sinti dialects and the presence of a relevant German lexical layer in Sinti is a secondary but important diagnostic feature of the group. The interference in some cases may also involve other dimensions of the dialects, such as prosody and phonetics.

<sup>16</sup> Caccini used the Italian writing system for his texts in Shinto Rosengro. Since not every sequence may be transcribed unequivocally into the writing system used in Romani scientific literature, I chose to maintain the system of the source.

<sup>17</sup> In the corpus of Shinto Rosengro, the pluperfect shows the person marker in one occurrence only.

<sup>18</sup> As previously mentioned about the example of Aija Varvara Romani, in Balkan dialects (together with Central and North-eastern dialects), the third person of the perfect of intransitive verbs is a participial form, which concords with the subject in number and, as far as the singular concerns, in gender; e.g. *gelo* ‘he went’, *geli* ‘she went’, *gele* ‘they went’. (cf. Boretzky & Iglá 2004, Teil 2: Karte 126).

<sup>19</sup> Formerly called Bangaru.

<sup>20</sup> As far as the Domari concerns, the example (7) showed that in Macalister 1914 two past forms, *āšti* and *āšta*, are recorded. Even though the origin of initial *a-* represents a problem, the endings *-i* and *-a* in *āšt-i* and *āšt-a* could be the outcomes of the endings of the verbal adjectives OIA \**sthitakah* (masculine) and \**sthitikā* (feminine). However, in Domari, *-i* and *-a* are also attested as tense/aspect markers suffixed to the inflected past forms (e.g. *kardom* ‘I did’ vs. *kardom-i* ‘I have done’ vs. *kardom-a* ‘I had done’, cf. Matras 2012: 11), and that makes difficult to establish with certainty their origin in *āšti* and *āšta*.

## REFERENCES

Beníšek 2020: Beníšek, M. The historical origins of Romani. In: Matras, Y., Tenser, A. (Eds.): *The Palgrave Handbook of Romani Language and Linguistics*. Cham: Palgrave Macmillan, 13-47.

Benveniste 1946: Benveniste, É. Structures des relations de personne dans le verbe. – *Bulletin de la Société de Linguistique*, 43, 1, 1-12.

Bloch 1932: Bloch, J. Survivance de skr. *āsīt* en indien moderne – *Bulletin de la Société de linguistique de Paris*, 33, 1, 55-65.

Boretzky, 2000: Boretzky, N. South Balkan II as a Romani dialect branch: Bugurdži, Drindari, and Kalajdži. – *Romani Studies*, fifth series, 10, 105-183.

Boretzky 1998: Boretzky, N. Erli. Eine Bestandsausnahme nach den Texte von Gilliat-Smith. – *Studii Romani*, 5-6, 122-160.

Boretzky 1996: Boretzky, N. Arli. Materialien zu einem südbalkanischen Romani-Dialekt. – *Grazer Linguistische Studien*, 46, 1-31.

Boretzky 1995: Boretzky, N. Die Entwicklung der Kopula im Romani. – *Grazer Linguistische Studien*, 43, 1-50.

Boretzky 1994: Boretzky, N. *Romani. Grammatik des Kalderaš-Dialekts mit Texten und Glossar*. Wiesbaden: Harrassowitz.

Boretzky, Cech, Iгла 2008: Boretzky, N., Cech, P., Iгла, B. *Die Südbalkanischen Dialekte und ihre innere Gliederung. Analyse und Karten*. Graz: Institut für Sprachwissenschaft.

Boretzky & Iгла 2004: Boretzky, N. & Iгла, B. *Kommentierter Dialektatlas des Romani. Vergleich der Dialekte*. Wiesbaden: Harrassowitz.

Bühler 1934: Bühler, K. *Sprachtheorie*. Stuttgart: Fischer.

Bubenik 1995: Bubenik, V. On typological changes and structural borrowing in the history of European Romani. In: Matras, Y. (ed.): *Romani in Contact. The History, Structure and Sociology of a Language*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1-24.

Caccini, Barontini, Piasere 2001: Caccini, S., Barontini, M., Piasere, L. *La lingua degli Shinte rosengre e altri scritti*. Roma: CISU.

Elšík & Beníšek 2020: Elšík, V., Beníšek, M. Romani dialectology. In: Matras, Y., Tenser, A. (Eds.): *The Palgrave Handbook of Romani Language and Linguistics*. Cham: Palgrave Macmillan, 389-428.

Elšík & Matras 2006: Elšík, V., Matras, Y. *Markedness and language change: The Romani sample*. Berlin: Mouton de Gruyter.

Gjerdman & Ljungberg 1963: Gjerdman, O., Ljungberg, E. *The Language of the Swedish Coppersmith Gipsy Johan Dimitri Taikon. Grammar, Texts, Vocabulary and English Word-index. Acta Academiae Regiae Gustavi Adolphi XL*. Uppsala/Copenhagen: Lundequistska Bokhandeln/Ejnar Munksgaard.

Grierson 1916: Grierson, G. A. *Linguistic survey of India*, Vol. IX, Part I. Calcutta: Superintendent Government Printing.

Hancock 1995: Hancock, I. *A Handbook of Vlax Romani*. Columbus: Slavica Publishers, Inc.

Herin 2012: Herin, B. The Domari language of Aleppo (Syria). – *Linguistic Discovery*, 10, 2, 1-52.

Iгла 1996: Iгла, B. *Das Romani von Ajia Varvara*. Wiesbaden: Harrassowitz.

Macalister 1914: Macalister, R. A. S. *The Language of the Nawar of Zutt, the Nomad Smiths of Palestine*. London: Bernard Quaritch. [Gypsy Lore Society Monographs 3].

Masica 1991: Masica, C. P. *The Indo-Aryan Languages*. Cambridge: Cambridge University Press.

Matras 2012: Matras, Y. *A Grammar of Domari*. Berlin/Boston: Walter de Gruyter.

- Matras 2005: Matras, Y. The classification of Romani dialects: A geographic-historical perspective. In: Ambrosch, G., Halwachs, D. and Schrammel, B. (Eds.): *General and Applied Romani Linguistics*. Muenchen: Lincom Europa, 7-26.
- Matras 2002: Matras, Y. *Romani. A Linguistic Introduction*. Cambridge: Cambridge University Press.
- Matras 2001: Matras, Y. Tense, aspect, and modality categories in Romani. – *STUF. Language Typology and Universals (Sprachtypologie und Universalienforschung)*, 53, 4, 162-180.
- Matras 1995: Matras, Y. (Ed.), *Romani in contact. The History, Structure and Sociology of a Language*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Meli 2019: Meli, G. *Il dialetto degli Shinte Rosengre: esame delle fonti e analisi della morfologia tra sincronia e diacronia*. (Phd Thesis, Università degli Studi di Milano)
- Niemen 1995: Niemen, A. *O ker kun le penijá. La casa con le ruote*. Roma: Sinno Editrice.
- Oberlies 2005: Oberlies, T. *A historical grammar of Hindi*. Graz: Leykam.
- Oliviero 1978: Oliviero. Quanto i u kurape. – *Rom. In cammino*, numero unico, 9.
- Oliviero 1980: Oliviero. Da u ghinape du puri: “U vela indren dar vale giala vrin du vudar”. – *Rom. In cammino*, numero unico, 8.
- Oliviero 1981: Oliviero. Un racconto antico: “I perla più bari du veltu”. – *Rom. In cammino*, numero unico, 3.
- Paspati 1870: Paspati, A. G. *Études sur les Tchinghianés ou Bohémiens de l’Empire Ottoman*. Constantinople: Koroméla.
- Patkanoff 1908: Patkanoff, K.P. Some words on the dialects of the Transcaucasian Gypsies. – *Journal of the Gypsy Lore Society*, new series, 2, 246-266.
- Piasere 1996: Piasere, L. Sigismondo Caccini e gli Šinte rozengere. In: Piasere, L. (Ed.): *Italia Romani, volume secondo*. Roma: CISU, 119-175.
- Plungian & van der Auwera 2006: Plungian, V.A., van der Auwera, J. Towards a typology of discontinuous past marking. – *STUF. Language Typology and Universals (Sprachtypologie und Universalienforschung)*, 59, 4, 317-49.
- Pott 1845: Pott, A. Über die Sprache der Zigeuner in Syrien. In: Hoefler, A. (Ed.): *Zeitschrift für die Wissenschaft der Sprache* 1, 175-186.
- Reichenbach 1947: Reichenbach, H. *Elements of Symbolic Logic*. New York: Macmillan.
- RMS Database: Romani Morpho-Syntax Database. <<https://romani.humanities.manchester.ac.uk/rms/>>, retrieved on 11.12.2022
- Sampson 1926: Sampson, J. *The Dialect of the Gypsies of Wales, Being the Older Form of British Romani Preserved in the Speech of the Clan of Abram Wood*. Oxford: Clarendon Press.
- Scala forthcoming: Scala, A. Presente lungo e presente breve nella romani: evidenze dialettali e funzionalità originaria di un’opposizione.
- Scala 2020: Scala, A. La flessione dell’imperfetto nella romani d’Abruzzo e la sua genesi. – *Archivio Glottologico Italiano*, 105(2), 223-43.

Scala 2014: Scala, A. The mixed language of the Armenian Boshha (Lomavren) and its inflectional morphology: some considerations in light of Armenian dialectal variation. – *AION – Annali del Dipartimento di Studi Letterari, Linguistici e Comparati, Sezione linguistica*, 3. 233-250.

Soravia 1977: Soravia, G. *Dialetti degli zingari italiani*. Pisa: Pacini Editore.

Spinelli 1995: Spinelli S. (Ed.), *Baxtalo drom*. Pescara: Tracce.

Spinelli 1996: Spinelli S. (Ed.), *Baxtalo drom*. Pescara: Tracce.

Tenser 2005: Tenser, A. *Lithuanian Romani*. Muenchen: Lincom Europa.

Whitney 1896: Whitney, W. D. *A Sanskrit Grammar: Including Both the Classical Language, and the Older Dialects, of Veda and Brahmana*. Leipzig: Breitkopf & Härtel.

### ABBREVIATIONS

F feminine

M masculine

MIA Middle Indo-Aryan

NIA New Indo-Aryan

OIA Old Indo-Aryan

PFV perfective

PL plural

REM remoteness

SG singular

### **PhD Giulia Meli**

Independent scholar (PhD at Università degli Studi di Milano)

Milano, Italy

e-mail: giuliameli301@gmail.com