



# How diaspora shaped Romani throughout the centuries: From India, to Europe, and to the Americas

Evangelia Adamou

## ► To cite this version:

Evangelia Adamou. How diaspora shaped Romani throughout the centuries: From India, to Europe, and to the Americas. Foundation for endangered languages conference XXV, Dec 2021, Tirana, Albania. halshs-03354759

**HAL Id: halshs-03354759**

**<https://shs.hal.science/halshs-03354759>**

Submitted on 26 Sep 2021

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

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

# **How diaspora shaped Romani throughout the centuries: From India, to Europe, and to the Americas**

Evangelia Adamou  
French National Centre for Scientific Research, CNRS

## **Abstract**

Romani is a diaspora language *par excellence* as it is an Indic language spoken exclusively outside the Indian subcontinent, namely in Europe, the Americas, and Australia. Over the centuries, some Romani varieties ceased to be transmitted to younger generations (mainly in northern and western Europe), while others were maintained (mainly in the Balkans, central and eastern Europe, and in the Americas). All, however, were shaped by extensive multilingual practices with various European languages. In the first part of the article, I present an overview of the contact-induced changes noted in different Romani varieties. In the second part, I discuss dispositions as expressed by Roma from Romania and Colombia in two surveys. In both surveys, respondents are in favor of the intergenerational transmission of Romani at home along with the language most relevant in work-related interactions. In sum, Romani offers a unique opportunity to observe how multilingualism allowed Romani to be maintained in some diaspora settings, while changing to converge with the contact languages, but also how in other settings multilingualism led to language shift.

## **Introduction**

In this paper, I discuss the case of Romani diaspora communities, engaged in mobile practices throughout the centuries, to illustrate how language is maintained, lost, and changes in different settings. In the first part of the article, I offer an overview of how Romani was shaped by extensive multilingual practices throughout the centuries. In the second part, I discuss dispositions that promote multilingualism and intergenerational transmission of both Romani and the language most relevant in work-related interactions.

## **Romani: a diaspora language shaped by language contact**

Romani is an Indic language of the Indo-European stock and the only Indic language to be exclusively spoken outside of the Indian subcontinent. Nowadays, Roma are the largest ethnic minority in Europe: the European Commission estimates that 6 million Roma reside in the European Union, but this number doubles when considering Europe as a whole.<sup>1</sup> Twenty years ago, Bakker (2001) estimated that approximately 4.6 million Roma speak Romani in Europe, with countries like Albania and Belarus exhibiting the highest rates of Romani speakers among the local Romani population (95%) and countries like the United Kingdom and Spain exhibiting the lowest rates of L1-Romani speakers (0.5%-1% excluding recent migrants). In addition, there are somewhere between 1 and 4 million Roma living in the Americas, but the number of speakers is much lower (Bernal, 2003).

## **Romani: a diaspora Indic language with a European typological profile**

In the early 20th century, comparison of phonological changes in Romani and other Indic languages allowed Ralph Turner to estimate that Romani emerged in central India, a view that is currently accepted in Romani linguistics (Matras, 2002). Historical phonology further allowed to conclude that the ancestors of today's Romani speakers most likely moved

---

<sup>1</sup> [https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combating-discrimination/roma-and-eu/roma-integration-eu\\_en](https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combating-discrimination/roma-and-eu/roma-integration-eu_en)

to northwest India before the 4th century CE, that they remained in this area until the 8th or 9th centuries CE and that they arrived at the Byzantine Empire around the 10th century CE. Roma were typically engaged in trade-related activities that require the use of a variety of languages spoken in the markets, depending on time and place. Already in the 19th century, August Pott and Franz Miklosich retrace Roma's possible itinerary from India thanks to the study of borrowings. We still find in many contemporary Romani varieties several borrowings from Persian or/and Kurdish (e.g., *zor* 'strength', *baxt* 'luck'), Armenian (e.g., *grast* 'horse', *pativ* 'honor'), and Greek (e.g., numerals like *efta* 'seven', *oxto* 'eight', *enja* 'nine', and kinship terms like *papo/papus* 'grandfather') (Matras, 2002).

In addition to the numerous lexical borrowings, Romani grammar has been greatly shaped by contact with non-Indic languages. For example, Romani exhibits nominative-accusative alignment for full noun phrases, that is the S(ingle) argument of an intransitive clause and the A(gent)-like argument of a monotransitive clause are marked with the same case whereas the P(atient)-like argument of a monotransitive clause takes a different case which is more marked than the nominative case (for a cross-linguistic overview see Comrie, 2013). The same nominative-accusative marking is also found for the pronouns. In comparison, morphological ergative alignment is found in many Indic languages.<sup>2</sup> According to Butt (2001), Sanskrit and Hindi were patient-oriented languages and, despite the novel case morphology, the modern system can be considered as the continuation of the elder stages. Ergative alignment can also be reconstructed for Proto-Romani, i.e., Romani as spoken before contact with the European languages (Bubeník, 2000).

Another good example to illustrate contact-induced changes is that, through contact with Greek, Romani developed a definite article based on the Romani demonstratives. Figure 1 shows that most languages in Europe have definite articles or affixes, but that definite articles are absent in most languages spoken in India (Dryer, 2013a).

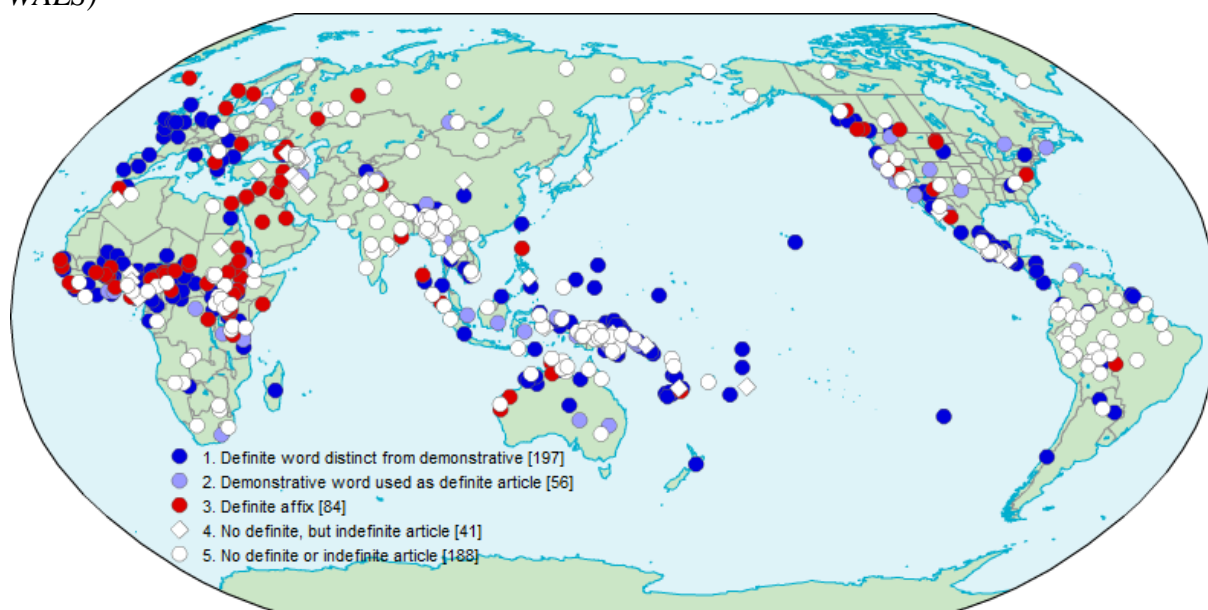
A third example concerns word order. At present, most Romani varieties have a dominant verb-object order (VO) whereas an object-verb order (OV) is encountered in most Indic languages, as can be seen in Figure 2 (Dryer, 2013b). The VO order in Romani is believed to be a change from a previously dominant verb-final order (Matras, 2002). VO is also the most frequent order in European languages as shown in Figure 2.

---

<sup>2</sup> In Comrie (2013), Hindi is not classified as an ergative language, but as having a tripartite alignment system, where S, A, and P are marked differently with some TMA specifications.

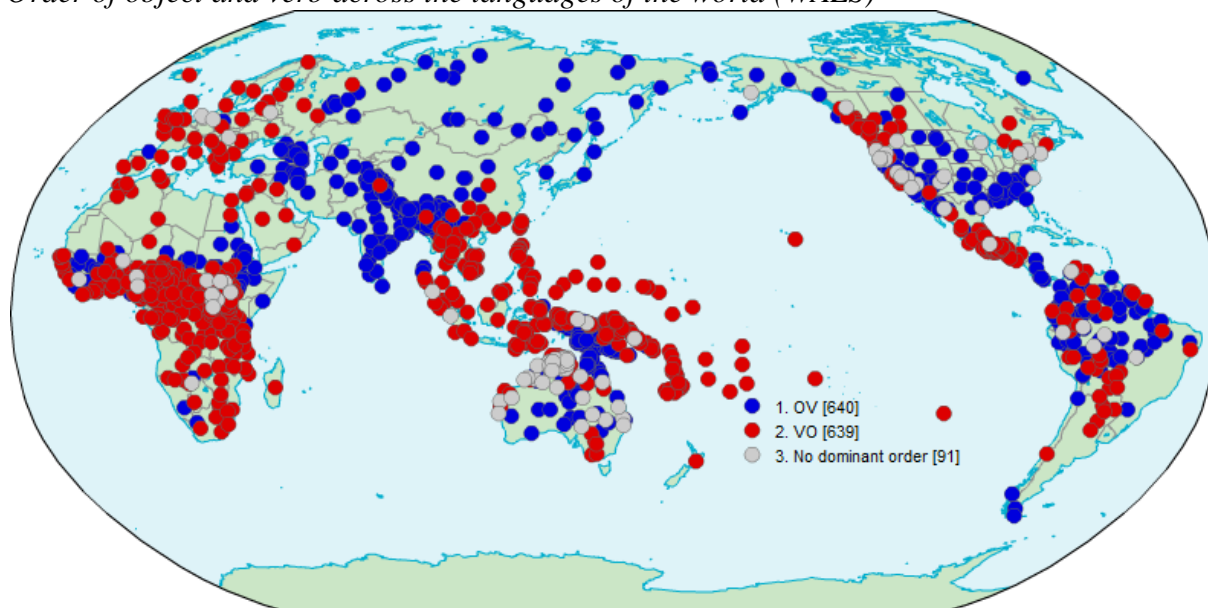
**Figure 1**

*Definite articles across the languages of the world (World Atlas of Language Structures, WALS)*



**Figure 2**

*Order of object and verb across the languages of the world (WALS)*



### **Romani contact-induced changes in different language contact settings**

Following the Ottoman expansion in the Balkans, several Romani groups fled the wars and migrated toward western and northern Europe (in the 14th-15th centuries). Over the centuries, some Romani varieties ceased to be transmitted to younger generations even if a part of the Romani lexicon is still in use (this is the case in the Scandinavian countries, the United Kingdom, Spain, and Portugal where revitalization efforts are ongoing), while others were maintained (mainly in the Balkans and central and eastern Europe). All, however, were shaped by extensive multilingual practices with the various European languages (Matras & Adamou, 2020).

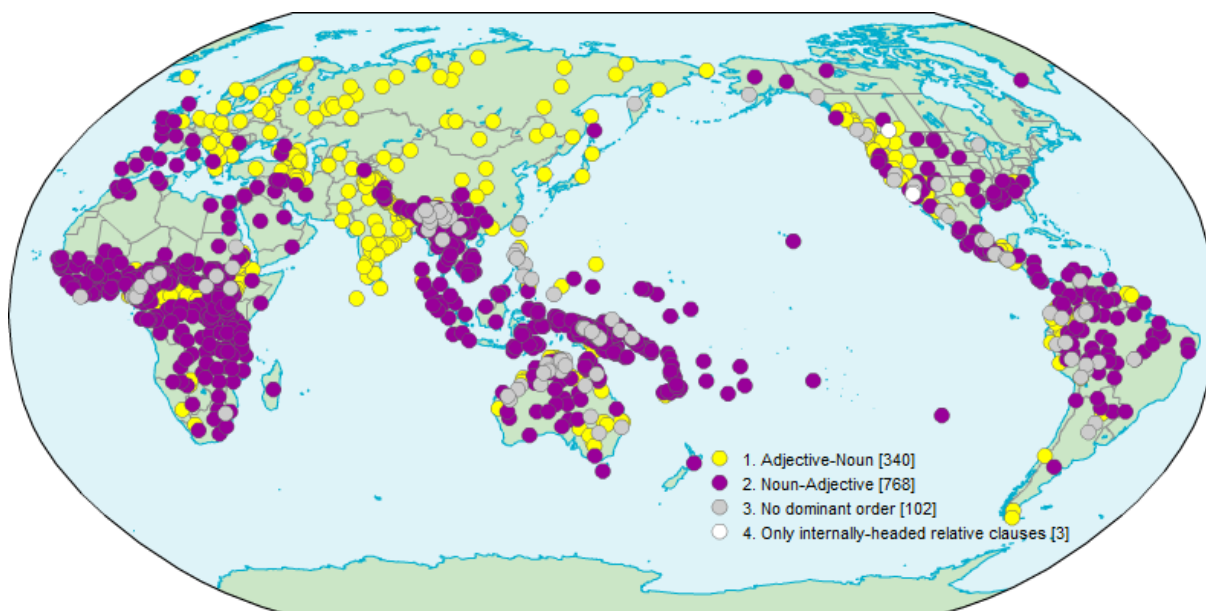
For example, the definite article was reduced or lost only in those varieties whose speakers were in long-term contact with languages that lack definite articles (Adamou & Matras, 2020). This is for example the case in the variety spoken by Polska Roma in Poland, in Lovari spoken in Russia, and in Finnish Romani spoken in Finland: all are varieties in long-term contact with languages that lack definite articles such as Polish, Russian, and Finnish.

Another example comes from the order of the adjective and the noun. In Romani, the adjective typically precedes the noun (AdjN). As can be seen in Figure 3, the AdjN order is a widespread feature among the languages of India and more generally of Eurasia apart from western Europe (Dryer, 2013c). Interestingly, Romani speakers who also speak a language with a dominant noun-adjective order (e.g., Romance languages like Romanian, Italian, and Spanish) are increasingly adopting a noun-adjective order (Adamou et al., 2021). In contrast, Romani speakers who also speak a language with a dominant adjective-noun order (e.g., Slavic languages) maintain this order in Romani too.

In addition to these contact-induced structural changes, codeswitching and borrowing are very frequent in most Romani varieties. For example, in a sample of 41 languages from across the world, Romani appears to be the highest borrower (Elšík, 2009). Moreover, some Romani varieties are well-known for exhibiting typologically-rare adaptation strategies. This is the case of Romani spoken in Greece and in Finland, two varieties that display a rare type of codeswitching/borrowing where L2-verbs are systematically used together with L2-morphology in otherwise Romani-dominant discourse, giving rise to so-called “unevenly mixed languages” (Adamou & Granqvist, 2015).

**Figure 3**

*Order of adjective and noun across the languages of the world (WALS)*



### **Romani in Latin America: the emergence of a new diaspora variety**

Beyond Europe, Romani is also the language of several million people living in the Americas. The presence of Roma in the Americas is documented as early as the colonial period (15th-18th centuries). Arrivals of newcomers resumed during the 19th century and have been continuous ever since. At present, Romani families live in practically every country of the American continent.

In the Latin American context, we note that multilingual practices with Spanish are giving rise to a characteristic diaspora Romani variety. For example, Mexican Romani has replicated the *ser* vs *estar* ‘to be’ distinction from Spanish (Adamou, 2013; Padure et al., 2018; Adamou et al., 2019). Specifically, Mexican Romani speakers have developed a distinction between attributive predications using the copula *si* ‘to be’ and the third person subject clitic pronouns in *l-*. In contrast, Romani speakers from all European varieties only use the copula *si* ‘to be’ in attributive predications. This is not a feature limited to Mexican Romani speakers and has since been noted in Colombia and Argentina. In Acuña, Adamou, and Sutre (in press), we adopt a cross-disciplinary approach (linguistic, historical, and ethnographic) that supports a scenario whereby this innovation spread throughout the various Romani groups via traditional, translocal social networks and mobile professional activities.

Another widespread influence of Spanish on Romani phonology, concerns the inherited voiced labial /v/ which is replaced by the Spanish bilabial approximant /β/, as can be seen in the inherited part of its vocabulary, for example, in words such as *βudar* ‘door’ and *βas* ‘hand’ (Adamou, 2013).

### **Romani dispositions towards language policy**

Official recognition of Romani as a minority language has been rare in the past, with the earliest and most noteworthy efforts in the Soviet Union from 1928 to 1936. Over the last forty years, however, this situation has changed and numerous European states, transnational institutions like the Council of Europe, foundations, and associations became engaged into the language policy and planning of Romani (Halwachs, 2020). In the Americas, this is the case only in Colombia. However, the contemporary understanding of “language policy” goes beyond state-controlled policy and includes the ways in which individuals and social groups act upon language policy through their practices (Spolsky, 2004). Critical scientists also call for language policy to be understood as a practice that aims to help speakers achieve specific cultural, educational, and economic goals (Canagarajah, 2008; Pennycook, 2010). Similarly, in a folk linguistics perspective, Albury (2014) proposes to take into consideration the knowledge and dispositions that individual speakers may have towards language policy.

During the Covid-19 pandemic, we conducted two online surveys to explore dispositions towards Romani language policy in two geographically distant settings: Romania and Colombia (Padure & Adamou, submitted; Acuña Cabanzo & Adamou, unpublished). These two settings share some similarities regarding institutional language policy, but also exhibit notable differences. Indeed, both countries officially recognize Romani as a minority language, but Romania incorporated the Romani language in the education system in the 1990s whereas in Colombia such measures are recent and considerably less developed. Also, population numbers differ considerably with estimates of Roma in Romania varying between 1.2 and 2.5 million and in Colombia between 2,500 and 5,000 people. In Colombia, Romani is overall learned by the younger generations in greater proportions than among other migrant or Indigenous groups: the census mentions 76.9% of Romani speakers, the highest rate for any minority language in the country (Paternina Espinosa, 2013). In Romania, estimates vary widely between 500,000 and 2.5 million speakers of Romani.

Despite present and historical differences that characterize Roma in the two settings, studies show that many share similar types of social relations, where individuals are tightly connected through family networks, and perhaps more loosely connected through social networks such as churches and schools (Toma & Foszto, 2018; Acuña Cabanzo, 2019). In addition, in both settings, production and labour processes share some similarities: employment can be local, national or transnational, within Europe for Romanian Roma and across the American continent for Colombian Roma (Sutre 2017; Acuña Cabanzo, 2019). Such mobilities are not only due to economic factors, but also to adverse situations that

cannot be adequately controlled by the state, the police, and the military. Indeed, Roma in both settings experience violence in daily life, whether crime-related, through drug trafficking for example, or due to civil conflict as in the case of Colombia, or to structural violence stemming from poverty, exclusion, and discrimination. In terms of employment the situation varies, but for many Roma, employment entails speaking languages other than Romani. As a result, daily life in the two settings may show some parallels: family relations are at the forefront, promoting the maintenance of Romani, while relations with outsiders are crucial at work, promoting language interactions in a language other than Romani. Finally, access to technologies helps sustain translocal communication networks through texting and social media (Acuña Cabanzo, 2019).

Following Albury (2014), we designed two online surveys with statements of knowledge or dispositions (i.e., statements that included evaluation and an attitudinal stance). Respondents were asked to indicate on a five-point scale their agreement with various statements or to choose among several responses. The statements from the original study were adapted to the specificities of the Romanian and Colombian contexts respectively in a two-step process: first, they were modified according to our understanding of the specific contexts; and second, they were discussed with Romani collaborators and modified according to suggested formulations. 180 respondents from Romania and 28 from Colombia completed the survey. In the Romanian survey, 86.5% of respondents declared residing in Romania, 80.5% declared being Roma, but only 65% declared speaking Romani. In the Colombian survey, 87% declared residing in Colombia, 96% self-identified as Roma, and 86% of respondents declared speaking Romani.

A comparison of responses in the two surveys shows that the dispositions expressed present some commonalities, in particular, with respect to the importance of multilingual practices: bilingual child acquisition at home is favoured by both groups (96% of Romanians and 89% of Colombians); parental transmission of Romani is seen as key (82% of Romanian respondents and 93% of Colombian respondents); general multilingualism is the preferred language model as opposed to monolingualism (66% of Romanian and 85% of Colombian respondents (strongly) disagreed with the statement ‘It would be better if everyone in the country spoke one language in all situations’); Romani language is viewed as an intrinsic part of Romani culture (85% of Romanian respondents and 75% of Colombian respondents noted that Romani is part of Romani culture), even though only 30% of Romanian and 46% of Colombian respondents considered that one has to speak Romani to be ‘a real Rom’; speaking Romani is primarily associated with interactions with family and friends rather than in formal settings (in reaction to the statement ‘In revitalising the language, people who can speak Romani should do so...’, 85-89% of respondents in both groups suggested that they should speak Romani ‘when talking with family and friends’ but only 32-34% of respondents in both groups declared that Romani should be used at work. Moreover, 40% of Romanian and 21% of Colombian respondents declared that Romani should also be used in official situations such as in parliament or in dealings with government officials).

In addition, responses to the two surveys shed new light on the impact that institutional practices have on Romani language ideologies in the long term. For example, we found that respondents from Colombia consider that Romani needs to be mainly spoken rather than written, whereas respondents from Romania consider both speaking and reading/writing to be crucial to support Romani. Although respondents from Colombia declared that education could play a role in Romani maintenance, they considered families and communities to be the primary agents. In contrast, respondents from Romania noted that education plays a central role in the transmission of Romani, in addition to the role played by families and

communities. Finally, respondents from Colombia were in favour of solutions that are less dependent on intervention from institutional actors, for example they expressed their support for the use of Spanish words in Romani rather than the wholesale creation of Romani words. In contrast, respondents from Romania opted for the creation of Romani words based on Indic roots, a process that requires the intervention of a learned stakeholder. The differences in the responses of the two groups regarding the role of education confirm observations on how institutional language arrangements shape individual dispositions (Padure & Adamou, submitted).

### Conclusion

In sum, in the first part of the paper, I discuss how in the absence of normative linguistic institutions, complex multilingual practices rapidly modify the grammar of a diaspora language such as Romani. In addition to borrowings, significant changes in the grammar are noted for Romani during Byzantine times, then in successive diaspora settings in Europe and in the Americas depending on the new contact languages. In the second part of the paper, the comparison of the dispositions expressed in two surveys, one with respondents from Colombia and one with respondents from Romania, shows that they overall value multilingualism and the transmission of Romani from one generation to the other. The comparison of responses further reveals that once institutional language policies are established, for example by including Romani in the education system as is the case in Romania, they promote a language ideology whereby language maintenance is ensured through schools in addition to families. Similarly, institutional language policy based on education endorses the significance of Romani in reading and writing in addition to oral interactions and promotes top-down lexical creation processes rather than everyday processes of lexical borrowing available to individual speakers and observed in most Romani varieties across Europe and the Americas up until today. To conclude, Romani offers a unique opportunity to observe how a language is maintained, lost, and changes throughout the centuries in successive diaspora settings.

### References

- Acuña Cabanzo, E. (2019). *Tracing the Romani Atlantic. An ethnography of trans-local interconnections and mobilities among Romani groups*. Doctoral dissertation at the University of Freiburg, Freiburg, Germany.
- Acuña Cabanzo, E., & Adamou, E. (unpublished). Survey on Romani disposition towards language policy in Colombia.
- Acuña Cabanzo E., Adamou, E., & Sutre, A. (in press). Contact-induced change and mobility: A cross-disciplinary approach to Romani in Latin America. *Language Ecology*.
- Adamou, E. (2013). Replicating Spanish *estar* in Mexican Romani. *Linguistics*, 51, 1075-1105. <https://doi.org/10.1515/ling-2013-0045>
- Adamou, E., De Pascale, S., García-Márkina, Y., & Padure, C. (2019). Do bilinguals generalize *estar* more than monolinguals and what is the role of conceptual transfer? *International Journal of Bilingualism*, 23, 1549-1580. <https://doi.org/10.1177/1367006918812175>
- Adamou E., Feltgen, Q., & Padure, C. (2021). A unified approach to the study of language contact: Cross-language priming and change in adjective/noun order. *International Journal of Bilingualism* online first <https://doi.org/10.1177/13670069211033909>
- Adamou, E., & Granqvist, K. (2015). Unevenly mixed Romani languages. *International Journal of Bilingualism* 19(5), 525-547.

- Adamou, E., & Matras, Y. (2020). Romani syntactic typology. In Y. Matras & A. Tenser (Eds.), *Handbook of Romani language and linguistics* (pp. 187-227). Palgrave Macmillan.
- Albury, N. (2014). "If we lose their language we lose our history": Knowledge and disposition in Māori language acquisition policy. *Journal of Language, Identity and Education*, 17, 69-84.
- Bakker, P. (2001). Romani in Europe. In G. Extra & D. Gorder (Eds.), *The other languages of Europe: Demographic, sociolinguistic and educational perspectives* (pp. 293-313). Multilingual Matters.
- Bernal, J. (2003). *The Rom in the Americas. Sub-commission on promotion and protection of human rights working group on minorities*. United Nations.
- Bubeník, V. (2000). Was Proto-Romani an ergative language? In V. Elšík & Y. Matras (Eds.), *Grammatical relations in Romani. The noun phrase* (pp. 205-227). John Benjamins.
- Butt, M. (2001). A reexamination of the accusative to ergative shift in Indo-Aryan. In M. Butt & T. H. King (Eds.), *Time over matter: Diachronic perspectives on morphosyntax* (pp. 105-141). CSLI Publications.
- Canagarajah, A. S. (2008). *Reclaiming the local in language policy and practice*. Routledge.
- Comrie, B. (2013). Alignment of case marking of full noun phrases. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/37>)
- Dryer, M. S. (2013a). Definite articles. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/37>)
- Dryer, M. S. (2013b). Order of object and verb. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/83>)
- Dryer, Matthew S. (2013c). Order of adjective and noun. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/87>)
- Elšík, V. (2009). Loanwords in Selice Romani, an Indo-Aryan language of Slovakia. In M. Haspelmath & U. Tadmor (Eds.), *Loanwords in the world's languages. A comparative handbook* (pp. 260-303). Mouton de Gruyter.
- Halwachs, D. W. (2020). Language policy and planning in Romani. In Y. Matras & A. Tenser (Eds.), *The Palgrave handbook of Romani language and linguistics* (pp. 429-458). Palgrave Macmillan.
- Matras, Y. (2002). *Romani: A linguistic introduction*. Cambridge University Press.
- Matras, Y., & Adamou, E. (2020). Romani and contact linguistics. In Y. Matras & A. Tenser (Eds.), *Handbook of Romani language and linguistics* (pp. 329-352). Palgrave Macmillan.
- Padure, C., & Adamou, E. (submitted). Cadrul sferelor de activitate aplicate politicii lingvistice: Dispoziții cu privire la revigorarea limbii rromani în România.
- Padure, C., De Pascale, S., & Adamou, E. (2018). Variation between the copula *si* "to be" and the *l*-clitics in Romani spoken in Mexico. *Romani Studies* 28: 263-292.
- Paternina Espinosa, H. A. (2013). *El Proceso organizativo del pueblo Rrom (Gitano) de Colombia (PRORROM). De la auto-invisibilidad como estrategia de resistencia étnica y cultural, a la visibilización como mecanismo de reconocimiento de derechos económicos, sociales, políticos y culturales*. Doctoral dissertation at the Universidad Autónoma de Madrid, Madrid, Spain.
- Pennycook, A. (2010). *Language as a local practice*. Routledge.
- Spolsky, B. (2004). *Language policy*. Cambridge University Press.

Sutre, A. (2017). *Du parcours du monde à son invention. Géographies tsiganes en Amérique du Nord des années 1880 aux années 1950*. Doctoral dissertation at the Ecole des Hautes Etudes en Sciences Sociales, Paris, France.

Toma, S., & Foszto, L. (2018). Returnees and their neighbors: Migration of the Romanian Roma, networks, social distance, and local development. *Review of Sociology*, 28(4), 37-60.